UAB
Universitat Autònoma de Barcelona

Enhancing recombinant adeno-associated virus (rAAV) production through cell and gene engineering

Laia Bosch Molist

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: https://creativecommons.org/licenses/?lang=ca

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: https://creativecommons.org/licenses/?lang=es

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/?lang=en

Escola d'Enginyeria

Departament d'Enginyeria Química, Biològica i Ambiental

Programa de Doctorat en Biotecnologia

Enhancing recombinant adeno-associated virus (rAAV) production through cell and gene engineering

Laia Bosch Molist

PhD Thesis, May 2024

Advisors:

Dr. Francesc Gòdia Casablancas

Dr. Laura Cervera Gracia

Dr. Miguel García Martínez

Prof. Dr. Francesc Gòdia Casablancas, Dr. Laura Cervera Gracia and Dr. Miquel García Martínez

Hereby certify:

That Laia Bosch Molist has successfully completed his PhD entitled **"Enhancing recombinant adeno-associated virus (rAAV) production through cell and gene engineering**" in the Chemical, Biological and Environmental Engineering Department at Universitat Autònoma de Barcelona with our supervision. The PhD thesis is hereby presented in this document in order to opt to the PhD title as Doctor in Biotechnology by the Universitat Autònoma de Barcelona.

Thereby and for the appropriate purposes, we sign this document in Bellaterra, May 2024.

Dr. Francesc Gòdia Casablancas	Dr. Laura Cervera Gràcia	Dr. Miquel García Martínez
(co-director)	(co-director)	(co-director)

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: https://creativecommons.org/licenses/by-nc-nd/4.0/

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: https://creativecommons.org/licenses/by-nc-nd/4.0/

WARNING. The access to the contents of this doctoral thesis is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/by-nc-nd/4.0/

CONTENT

ABBREVIATIONS **8**

SUMMARY **10**

RESUMEN **12**

RESUM **14**

INTRODUCTION **17**

1. The gene therapy journey **18**

1.1. Historical context 18

1.2. Current status of gene therapy 19

1.3. Viral vectors 22

2. Adeno-associated virus (AAV) **24**

2.1. Adeno-associated virus genome and structure 25

2.2. Adenoviral helper functions for adeno-associated virus replication 30

2.3. Adeno-associated virus life cycle 34

3. Adeno-associated virus as a gene delivery vector **40**

3.1. Vectorization process 40

3.2. Need of cost-efficient rAAV manufacturing systems 43

4. Strategies for rAAV production **45**

4.1. Transient gene expression 47

4.2. Stable gene expression 54

5. Operation systems for rAAV production through triple transfection into suspension HEK293 cells **59**

5.1. Batch cultures 59

5.2. Fed-batch cultures 60

5.3. Continuous culture strategies: perfusion operation mode 60

5.4. The extended gene expression protocol 62

OBJECTIVES **63**

RESULTS **65**

PART I. Enhancement of rAAV production through transient transfection **67**

CHAPTER ONE 69

Quantitative proteomic analysis of HEK293 provides insights of host cell response associated to the extended gene expression method of rAAV production

CHAPTER TWO 103

Metabolic engineering of HEK293 cells to enhance rAAV production and export

CHAPTER THREE 137

Strategies to circumvent endogenous *E1A* gene expression bottleneck in HEK293 cells for enhanced rAAV production

PART II. Proof of concept of an inducible system towards stable rAAV cell line generation **171**

CHAPTER FOUR 173

Rational design of optimal inducible Rep construct for the generation of a stable rAAV packaging cell line in HEK293 cells

CHAPTER FIVE 209

Engineering an E1A-inducible A549 cell line for stable rAAV production

GENERAL DISCUSSION AND FUTURE DIRECTIONS **243**

CONCLUSIONS **256**

REFERENCES **259**

ACKNOWLEDGMENTS/ AGRAÏMENTS **291**

ABBREVIATIONS

AAV: Adeno-Associated Virus

AAVS1: Adeno-Associated Virus Site 1

AdV: Adenovirus

ATF: Alternating tangential flow

BHK: Baby Hamster Kidney

BV: Baculovirus

CDE: Cell Density Effect

CDS: Coding DNA Sequence

CHO: Chinese Hamster Ovary

CMV: Cytomegalovirus

cGMP: current Good Manufacturing Practice

DLS: Dynamic Light Scattering

DoE: Design of Experiments

Dox: Doxycycline

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic Acid

EGE: Extended Gene Expression

EMA: European Medicines Agency

ESCRT: Endosomal Sorting Complex Required for Transport

EV: Extracellular vesicle

F17: FreeStyle F17™ Cell Culture Media

FASP: Filter Aided Sample Preparation

FDA: Food and Drug Administration

FDR: False Discovery rate

GFP: Green Fluorescent Protein

HCD: High Cell Density

HEK293: Human Embryonic Kidney 293

HeLa: Henrietta Lacks

hbt: Hours Before-Transfection

hpi: Hours Post-Infection

hpt: Hours Post-Transfection

hptd: Hours Post-Transduction

HPV: Human papillomavirus

HMGCS1: Hydroxymethylglutaryl-CoA Synthase 1

ILV: Intraluminal vesicles

ITR: Inverted Terminal Repeat

LC-MS/MS: Liquid Chromatography coupled with tandem mass spectrometry

ME: Medium Exchange

MOI: Multiplicity of Infection

MS: Mass spectrometry

MVB: Multi Vesicular Bodies

NTA: Nanoparticle Tracking Analysis

nt: Nucleotides

ORF: Open Reading Frame

PBS: Phosphate-Buffered Saline

PCR: Polymerase Chain Reaction

PEI: Polyethylenimine

PTM: Post-Translational Modification

PVDF: Polyvinylidene fluoride

qPCR: Quantitative Polymerase Chain Reaction

rAAV: Recombinant Adeno-Associated Virus

rAdV: Recombinant Adenovirus

rBV: Recombinant Baculovirus

Rb: Retinoblastoma protein

rcAAV: Replication-competent AAV or wt AAV-like particles

RSM: Response Surface Methodology

RT: Re-Transfection

RT-qPCR: Reverse Transcription Quantitative Polymerase Chain Reaction

rtTA: Reverse Tetracycline-Controlled Transactivator

SDS: Sodium Dodecyl Sulphate

SGE: Stable Gene Expression

Sf9: *Spodoptera frugiperda* 9

shRNA: Short Hairpin RNA

SN: Supernatant

TFA: Trifluoroacetic acid

TEAB: Tetraethylammonium tetrahydroborate

TFF: Tangential Flow Filtration

TGE: Transient Gene Expression

TMT: Tandem Mass Tag

TT: Triple Transfection

TU: Transducing Units

UPR: Unfolded Protein Response

VCD: Viable Cell Density

WB: Western Blot

wt: wild-type

SUMMARY

Gene therapy represents a transformative therapeutical innovation, aiming to correct or counteract genetic mutations and their associated pathologies by introducing, modifying, or suppressing specific genes within an individual's cells. In recent years, recombinant adeno-associated viruses (rAAVs) have emerged as one of the most promising vectors for gene therapy applications and several rAAV-based therapies have been approved, particularly for the treatment of orphan diseases, increasing the demand of clinical-grade rAAV material. However, the complexity of rAAV production methods and the lack of robust scale-up current manufacturing processes, limits the capacity to produce rAAVs at the volumes required for clinical applications in a cost-effective manner. Consequently, the high cost of these products limits patient access to gene therapies, hindering the transition to treatments for prevalent diseases.

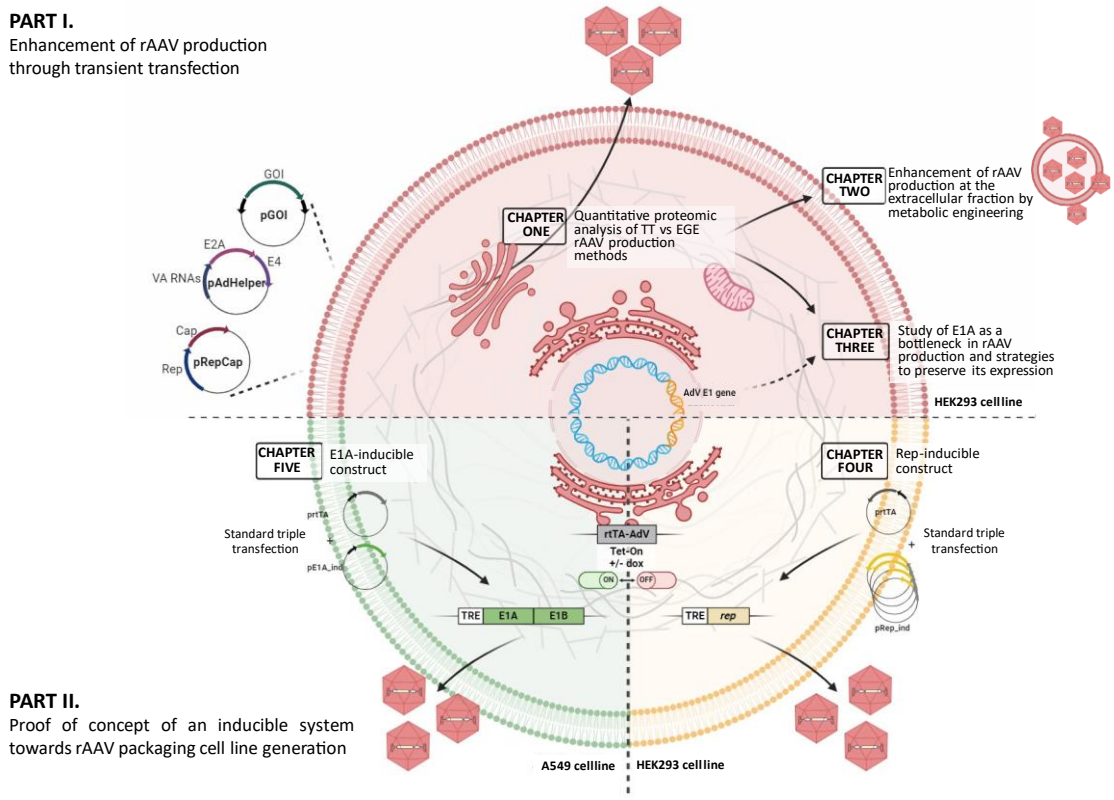
These challenges are addressed in this thesis by investigating the rAAV production bottlenecks and proposing cellular and genetic modifications to advance into the development of rAAV-based therapies, giving specific interest to rAAV9 and rAAV1. This is addressed in two main parts.

The first part, comprising **Chapters one, two** and **three**, focuses on the development of optimization strategies to improve rAAV production through transient gene expression (TGE). Although triple transfection (TT) into HEK293 cells is one of the dominant rAAV production platform nowadays, its robust application is limited, and different approaches are explored to improve it. In previous work, the extended gene expression (EGE) method was developed as a promising tool to enhance rAAV production, by enabling the continuous harvest of the secreted rAAV fraction, outperforming the TT, thereby eliminating the need of cell lysis. In **Chapter one**, the conventional TT method is compared to the EGE approach of rAAV production at a molecular level using a proteomic quantitative approach. The comparison of differential protein expression and enriched molecular processes of HEK293 cells between the two production methodologies provides insights into each cellular state, allowing to identify putative bottlenecks affecting the TT method of rAAV production and potential targets for rational engineering of host mammalian cell lines. These targets are used for host cell metabolic engineering in **Chapter two**, to enhance the rAAV harvest in the supernatant fraction. Additionally, media supplementation feeds are developed to mitigate the impact of cell density on transfection efficiency, which typically declines at higher cell concentrations limiting rAAV production by the TT method at large scale. In **Chapter three**, the dynamics in adenoviral E1A gene expression in HEK293 cells during rAAV production is studied

as a potential constrain for viral vector production. The E1A levels in a standard rAAV production system are characterized and different strategies to preserve its expression, thereby enhancing rAAV production, are proposed.

The second part, comprising **Chapters four** and **five**, focuses on developing strategies towards the proof of concept for a stable gene expression (SGE) method of rAAV production, evaluating different cell lines as hosts. Gene engineering techniques for inducible gene expression based on the Tet-On system are employed alongside the Design of Experiments (DoE) tool to develop and optimize rAAV production methods. In **Chapter four**, a first step towards the SGE method using an adenoviral E1-containing cell line such as HEK293 is explored, based on a stepwise engineering approach for inducible *rep* gene expression. In **Chapter five**, the shift is focused on a cell line lacking endogenous E1 expression, such as A549 cells, and the inducibility of E1A gene is evaluated, establishing A549 cells as a promising platform for rAAV production.

Overall, the results presented in this thesis contribute to the development of more efficient and cost-effective rAAV production strategies, using both transient gene expression and stable cell line production platforms.



Graphical abstract of the thesis. In Part I, the enhancement of rAAV production through transient transfection in HEK293 cells is addressed in **Chapter one, two** and **three**. In Part II, an inducible system towards stable rAAV cell line generation is evaluated, by inducible *rep* gene expression in HEK293 cell line (**Chapter four**) or *E1A* inducible gene expression in A549 cells (**Chapter five**).

RESUMEN

La terapia génica representa una revolución dentro de las terapias emergentes, con el objetivo de corregir o mitigar mutaciones genéticas y sus enfermedades relacionadas a través de la introducción, alteración o eliminación de genes específicos en las células de un individuo. Los virus adenoasociados recombinantes (rAAVs) han surgido como vectores altamente efectivos para esta finalidad, con varias terapias basadas en rAAVs ya aprobadas para tratar enfermedades raras, aumentando la demanda de rAAVs de calidad clínica. No obstante, la complejidad de los métodos de producción de rAAVs y la falta de procesos de fabricación robustos eleva los costos de las terapias génicas, restringiendo el acceso de los pacientes a estos tratamientos y dificultando su aplicación en enfermedades más comunes.

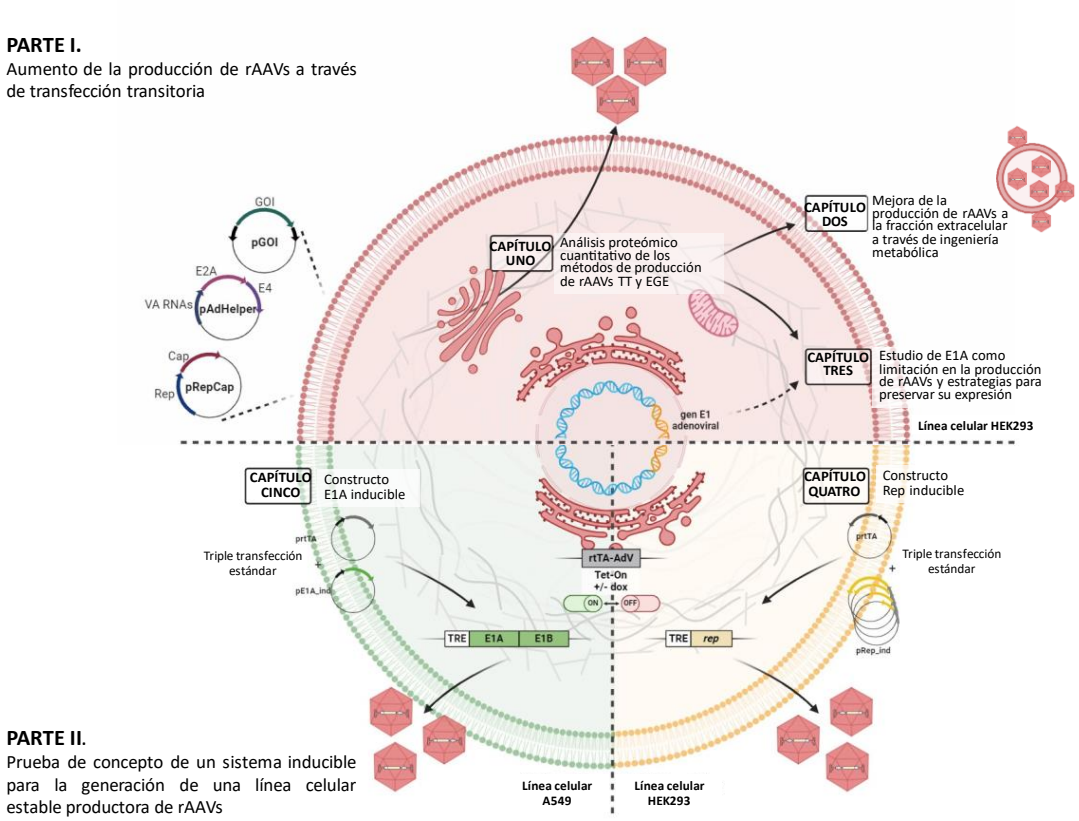
Para abordar estos desafíos, se investigan las limitaciones en la producción de rAAVs y se proponen modificaciones celulares y genéticas para avanzar en el desarrollo de terapias basadas en rAAVs, especialmente en rAAV9 y rAAV1. Este análisis se divide en dos partes principales.

La primera parte, que comprende los **Capítulos uno, dos** y **tres**, se centra en el desarrollo de estrategias de optimización para mejorar la producción de rAAVs mediante la expresión génica transitoria (TGE). Aunque actualmente la triple transfección (TT) en células HEK293 es una de las plataformas dominantes para la producción de rAAVs, su implementación presenta limitaciones significativas, por ese motivo, se exploran métodos alternativos. En trabajos anteriores, el método de expresión génica extendida (EGE) se desarrolló como una herramienta prometedora para mejorar la producción de rAAVs, permitiendo superar la producción obtenida por el método de TT, recogiendo los rAAVs del medio extracelular, eliminando la necesidad de lisis celular. En el **Capítulo uno**, el método de producción de rAAVs por TT se compara con el método de EGE a nivel molecular empleando proteómica cuantitativa. La comparación de la expresión proteica diferencial y los procesos moleculares alterados entre las dos metodologías de producción de las células HEK293, proporciona un análisis detallado sobre cada estado celular, permitiendo identificar posibles limitaciones que afectan el método de producción por TT y dibujando posibles objetivos para la ingeniería racional de líneas celulares. En el **Capítulo dos**, se utilizan las dianas identificadas para mejorar la obtención de rAAVs en la fracción extracelular a través de ingeniería metabólica en las células huésped. Además, se estudia la suplementación del medio para mitigar los efectos negativos de la alta densidad celular sobre la eficiencia de transfección. En el **Capítulo tres**, se estudia la expresión del gen adenoviral *E1A* durante la producción de rAAVs en células

HEK293 como una posible limitación para la producción de vectores virales. Se caracteriza la dinámica de expresión de los niveles de E1A durante la producción de rAAVs y se proponen estrategias para preservar su expresión, mejorando así la producción.

La segunda parte, que comprende los **Capítulos cuatro** y **cinco**, se centra en desarrollar estrategias hacia la prueba de concepto para un método de expresión génica estable (SGE) de producción de rAAVs, evaluando el uso de diferentes líneas celulares como sistemas productores. Para ello, se emplean técnicas de ingeniería génica para la expresión génica inducible basada en el sistema Tet-On, junto con la herramienta de Diseño de Experimentos (DoE). En el **Capítulo cuatro**, se explora la expresión del gen *rep* de forma inducible usando las células HEK293, que contienen el gen adenoviral *E1*. En el **Capítulo cinco**, el enfoque se centra en una línea celular que carece de expresión endógena de E1, como las células A549 y se evalúa la inducibilidad del gen E1A, estableciendo las células A549 como una plataforma prometedora para la producción de rAAVs.

En general, los resultados presentados en esta tesis contribuyen al desarrollo de estrategias de producción de rAAVs más eficientes y con menores costos, utilizando tanto plataformas de expresión génica transitoria como de líneas celulares estables.



Resumen gráfico de la tesis. En la Parte I, que comprende los **Capítulos uno, dos y tres**, se aborda la mejora de la producción de rAAV mediante la transfección transitoria en células HEK293. En la Parte II, se evalúa un sistema inducible para la generación de líneas celulares estables productoras de rAAVs, mediante la expresión inducible del gen *rep* en la línea celular HEK293 (**Capítulo cuatro**) o la expresión inducible del gen *E1A* en células A549 (**Capítulo cinco**).

RESUM

La teràpia gènica representa una revolució dins les teràpies emergents, amb l'objectiu de corregir o contrarestar mutacions genètiques i les seves malalties associades a través de la introducció, modificació o supressió de gens específics en les cèl·lules d'un individu. Els virus adenoassociats recombinants (rAAVs) han sorgit com a vectors altament efectius per a aquesta finalitat, amb diverses teràpies basades en rAAVs ja aprovades per tractar malalties rares, augmentant la seva demanda en qualitat clínica. No obstant això, la complexitat dels mètodes de producció de rAAVs i la falta de processos de fabricació robustos incrementen els costos de les teràpies gèniques, restringint l'accés dels pacients a aquests tractaments i dificultant la seva aplicació en malalties més comunes.

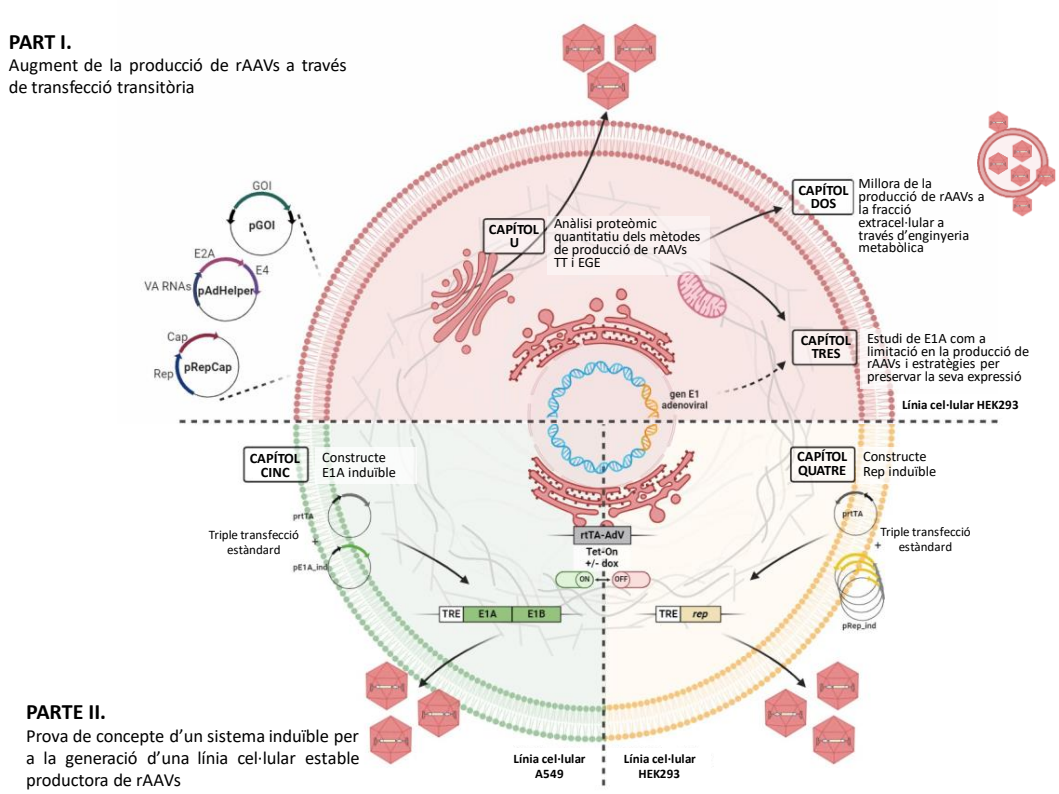
Per abordar aquests desafiaments, s'investiguen les limitacions en la producció de rAAVs i es proposen modificacions cel·lulars i genètiques per avançar en el desenvolupament de teràpies basades en rAAVs, especialment en rAAV9 i rAAV1. Aquesta anàlisi es divideix en dues parts principals.

La primera part, que comprèn els **Capítols u**, **dos** i **tres**, es centra en el desenvolupament d'estratègies d'optimització de la producció de rAAVs mitjançant l'expressió genètica transitòria (TGE). Encara que actualment la triple transfecció (TT) en cèl·lules HEK293 és una de les plataformes dominants per a la producció de rAAVs, la seva implementació presenta limitacions significatives, per aquest motiu, s'exploren mètodes alternatius per a la seva millora. En treballs anteriors, el mètode d'expressió genètica extesa (EGE) es va desenvolupar com una eina prometedora per millorar la producció de rAAVs, ja que permet superar la producció obtinguda pel mètode de TT, recollint els rAAVs en el medi extracel·lular, eliminant la necessitat de lisi cel·lular. Al **Capítol u**, el mètode de producció de rAAVs per TT es compara amb el mètode d'EGE a nivell molecular utilitzant proteòmica quantitativa. La comparació de l'expressió proteica diferencial i els processos moleculars de les cèl·lules HEK293 entre les dues metodologies de producció proporciona un anàlisi detallat sobre cada estat cel·lular, permetent identificar possibles limitacions que afectin el mètode TT de producció de rAAVs i possibles dianes per a l'enginyeria racional de línies cel·lulars. Al **Capítol dos**, s'utilitzen les dianes identificades per a l'enginyeria metabòlica de cèl·lules hoste, per millorar l'obtenció d'AAVs a la fracció extracel·lular. A més, s'estudia la suplementació del medi per mitigar els efectes negatius de l'alta densitat cel·lular sobre l'eficiència de transfecció. Al **Capítol tres**, s'estudia l'expressió del gen adenoviral E1A durant la producció de

rAAVs en cèl·lules HEK293 com una possible limitació per a la producció de vectors virals. Es caracteritza la dinàmica d'expressió dels nivells d'E1A durant la producció de rAAVs i es proposen estratègies per preservar la seva expressió, millorant així la producció.

La segona part, que comprèn els **Capítols quatre** i **cinc**, se centra en explorar estratègies per a la prova de concepte d'un sistema d'expressió genètica estable (SGE) de producció de rAAVs, avaluant l'ús de diferents línies cel·lulars com a sistemes productors. Per això, s'empren tècniques d'enginyeria genètica per a l'expressió induïble basada en el sistema Tet-On, juntament amb l'eina de Disseny d'Experiments (DoE) per desenvolupar i optimitzar els mètodes de producció de rAAVs. Al Capítol quatre, s'explora un primer pas cap al mètode SGE basat en l'expressió del gen *rep* de forma induïble utilitzant una línia cel·lular que conté el gen adenoviral *E1*, com la cèl·lula HEK293. El **Capítol cinc**, es centra en una línia cel·lular que marca l'expressió endògena d'E1, com les cèl·lules A549, i s'avalua la induïbilitat del gen *E1A*, establint les cèl·lules A549 com una plataforma prometedora per a la producció de rAAVs.

De forma global, els resultats presentats en aquesta tesi contribueixen al desenvolupament d'estratègies de producció de rAAVs més eficients i amb menors costos, utilitzant tant plataformes d'expressió genètica transitòria com de línies cel·lulars estables.



Resum gràfic de la tesis. La Part I, que comprèn els **Capítols u, dos i tres**, se centra en la millora de la producció de rAAVs mitjançant la transfecció transitòria en les cèl·lules HEK293. A la Part II, s'avalua un sistema induïble per a la generació de línies cel·lulars estables productores de rAAVs, mitjançant l'expressió induïble del gen *rep* en la línia cel·lular HEK293 (**Capítol quatre**) o l'expressió induïble del gen *E1A* en cèl·lules A549 (**Capítol cinc**).

INTRODUCTION

1. The gene therapy journey

Gene therapy is comprised by the European Medicines Agency (EMA) within the Advanced Therapy Medicinal Products (ATMPs). It is defined as an active substance containing or consisting of a recombinant nucleic acid used in or administered to human beings aiming for a therapeutic, prophylactic or diagnostic effect that relates directly to the recombinant nucleic acid sequence it contains, or to the product of genetic expression of this sequence[1]. Depending on the underlying genetic problem, there are different types of therapeutic approaches with a view of regulating, repairing, replacing, adding or deleting a genetic sequence. For example, if an endogenous gene is non-functional or missing, a functional therapeutic gene is added to supplement the deficient gene. In contrast, if a specific gene product is harmful or malfunctioning, the strategy is to delete or inactivate it, decreasing its expression levels by editing the incorrect gene sequences[2].

These approaches target the endogenous genetic and protein processing pathways associated with the etiology of hereditary diseases. Unlike conventional treatments, gene therapy has been envisioned as a potential one-time cure for many diseases[3].

When selecting a therapeutic approach, three crucial components must be considered: a) the type of vector used to deliver the therapeutic transgene into targeted cells, which can be either viral or non-viral; b) the nature of the delivered transgene into targeted cells, which may integrate into the genome or persist in an episomal form; c) the mode of administration, which can be either *in vivo*, where the gene is delivered directly into the patient's body, or *ex vivo*, where the target cells are removed from the patient, treated with the gene therapy and returned to the patient's body.

1.1. Historical context

The journey of gene therapy development was initiated with the identification of nucleic acid structure and the foundations of inheritance in the early 20th century[4] (**Figure 1**). Subsequent scientific advancements, such as virus-based methods for transforming mammalian cells in the 1960s and the development of recombinant DNA techniques in 1970s, laid the groundwork for gene manipulation, establishing viruses as promising tools to deliver genetic material into mammalian cells[5]. In 1972, Theodore Friedmann and Richard Roblin first introduced the concept of gene therapy[6]. They envisioned the use of genetic medicines to incorporate DNA sequences into patient cells to counteract genetic disorders, and emphasized the importance of establishing ethical and scientific criteria to guide gene therapy development[2,6].

This progress initiated the early age of gene therapy, with the first human gene therapy trial in 1990 for the treatment of two children with severe combined immunodeficiency (SCID) caused by a genetic defect in the adenosine deaminase (ADA) gene[7]. In this trial, a retroviral vector was used to deliver a normal copy of ADA gene into T cells collected from the same patient, modified outside the body (known as "*ex vivo*" gene therapy) and then reintroduced. Although optimal therapeutic levels for long term correction could not be achieved, it demonstrated the feasibility of gene therapy for monogenic diseases and identified basic strategies for further development[7,8].

This foundational study paved the way for the initiation of numerous gene transfer clinical trials for a wide range of hereditary diseases[9,10]. However, the field experienced a significant slowdown after the tragic death of Jesse Gelsinger in 1999. Gelsinger, afflicted with a partial deficiency of ornithine transcarbamylase (OTC), enrolled in a gene therapy clinical trial where he was treated with an adenoviral vector encoding a normal copy of the OTC gene. Tragically, he passed away four days post-treatment due to a severe immune reaction triggered by the high dose of adenovirus administered. This was the first patient experiencing serious adverse events directly linked to the viral vector treatment[11,12]. In response, the U.S Food and Drug Administration (FDA) and the National Institutes of Health (NIH) created new programs, namely the Gene Therapy Clinical Trial Monitoring Plan and Gene Transfer Safety Symposia to ensure safety and transparency of future gene therapy clinical trials[13]. Shortly after, in 2000, a clinical trial directed towards X-linked SCID gene therapy was initiated to replace the mutated interleukin-2 receptor gamma chain (IL2RG) gene using a retroviral vector[14]. Although it was successful, this vector contained an enhancer sequence that, following transgene insertion, activated an oncogene which led to T-cell proliferation. Consequently, 5 out of the 20 treated children developed leukemia[15].

Despite these circumstances slowed down the progress of gene therapy, they also highlighted the need for a deeper understanding of viral vector biology for gene transfer, transduction mechanisms, and the immune response to viral components. This knowledge is essential for the development of more specific and efficient gene transfer vectors[2].

1.2. Current status of gene therapy

The gene therapy field was able to overcome from this hurdle after several years, and since the start of the 21st century, several gene therapy-based products have been approved and numerous clinical trials have been successfully completed, expanding the field, and bringing gene therapies back into the spotlight. A key factor in this resurgence has been the development of safer viral

vectors, which have significantly enhanced the safety and efficacy of these therapies, renewing interest and confidence in their potential.

The first commercialized gene therapy product was Gendicine®, approved in China in 2003 for treating head and neck squamous cell carcinoma. In 2012, the EMA approved Glybera®, an adeno-associated virus (AAV)-based therapy for lipoprotein lipase deficiency (LPLD), which became the first gene therapy product approved in Europe[16]. Despite the initial breakthrough, Glybera was withdrawn from the market in 2017 due to its limited use[17].

Subsequent approvals followed. In 2016, Strimvelis® for ADA-SCID was approved by EMA as the first *ex-vivo* gene therapy product[18]. A year later, both FDA and EMA approved two *ex-vivo* treatments, Kymriah® and Yescarta®, in which T cells are targeted by lentivirus and retrovirus respectively, and genetically engineered to target CD19 receptors of cancerous B cells. The resulting chimeric T cell antigen receptor (CAR-T) cells are then reinfused into the patient. In 2017, Luxturna® treating a congenital form of retinal dystrophy (LCA), was the first *in-vivo* AAV-based gene therapy product approved by FDA[19]. After that, from 2019 to 2023 the approval of 9 more gene therapy products followed[20]. It is important to note that gene therapies, representing potential cures for unmet medical needs raise questions about patient access. Notably, Hemgenix®, consisting on an AAV vector encoding Factor IX for Hemophilia B bleeding disorder, has a price of $3.5M per dose[21]. Another example is Casgevy®, the first gene editing-based therapy approved by the FDA in 2023 for sickle cell disease, based on CRISPR/Cas9, which is priced at $2.2M[22].

This seven-figure price tags associated with gene therapies pose significant barriers on patient access. Patients who could benefit from these therapies may not be able to afford them, and the costs may be unsustainable for public health systems or insurance providers[23]. The small number of patients needing treatment and the complexity of the manufacturing systems, leading to high costs, are key factors behind this situation. Therefore, it is essential to develop more efficient production systems that can significantly reduce manufacturing costs and treatment prices, making these therapies more accessible.

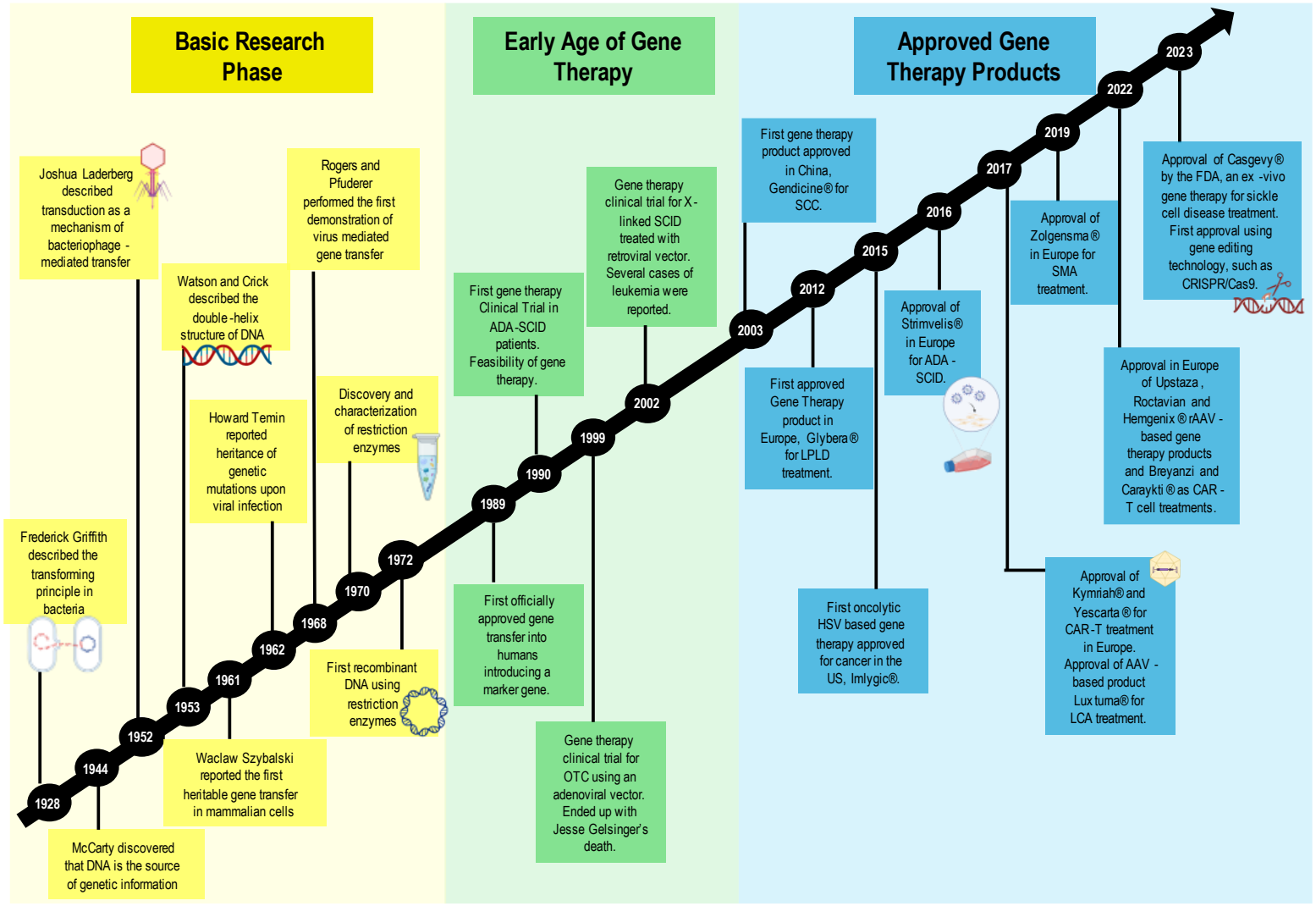


Figure 1. Timeline and key milestones in gene therapy development. Adapted from Wirth *et al.*[24] and Arabi *et al.*[2]. Information on approved products was obtained from the Paul-Ehrlich-Institut as of March 2024[20]. Illustrations were created with BioRender.com. ADA-SCID, Adenosine Deaminase Deficiency-Severe Combined Immunodeficiency; LCA, Leber Congenital Amaurosis; LPLD, Lipoprotein Lipase Deficiency; OTC, Ornithine Transcarbamylase Deficiency; SCC, Squamous Cell Carcinoma; SMA, Spinal Muscular Atrophy.

As of October 2023, 16 gene and cell therapies have been granted approval as ATMPs in Europe, a sign of the momentum that gene therapy has achieved in the last few years[25]. Moreover, the scope of gene therapy-based clinical trials has broadened, both in terms of the number of trials and the range of medical conditions they target. To date, more than 2500 gene therapy clinical trials are registered on clinicaltrials.gov, targeting a variety of diseases. The majority of which address cancer (65%), followed by monogenic diseases (11%) and infectious diseases (7%)[26].

Importantly, all the mentioned therapeutical products rely on the delivery vehicle to transport the payload to target cells and perform the therapeutic effect.

1.3. Viral vectors

For a gene therapy to be effective, the delivery vehicle must possess specific key characteristics: it must target the intended cell type specifically, transduce it with high efficiency, deliver the genetic material, and induce stable, long-term expression of the therapeutic gene at appropriate levels. Addressing the latter requires vector engineering techniques, which are disease-dependent and rely on the correct design of the therapeutic transgene within the vehicle. The former depends on the vector of choice. Different vectors are used for gene therapy with different characteristics. These can be classified into viral or non-viral vectors[27].

Non-viral vectors include nucleic acids covered by lipid nanoparticles (LNPs), exosomes, cationic polymers, inorganic nanoparticles and polymer hydrogels to enhance its stability[28]. LNPs are nowadays the most popular after its use in several COVID vaccines, based on the delivery of mRNA sequences encoding the SARS-COV-2 spike protein to the vaccine recipient cells as a defense against the SARS-COV-2, the virus responsible for the COVID-19 pandemic[28]. However, non-viral vectors often face challenges regarding effectiveness for long-term expression and cell-specificity *in vivo*, which may restrict its use in gene therapy clinical settings[29].

In contrast, viral vectors leverage the natural capacity of viruses to infect cells, a capability refined over thousands of years of evolution to evade host immune responses and efficiently transduce target cells. They are considered the most effective delivery vehicles, preferred in 70% clinical trials compared to non-viral counterparts which account for 30%[26]. Viral vectors are engineered to lack wild-type (wt) viral genes responsible for causing disease; during the vectorization process, viral regions necessary for replication are removed, while sequences essential for genome packaging and target cell transduction are maintained, delivering nucleic acids into cells via natural infection pathways[30].

They are further categorized based on their capacity for integration into the host genome into integrating and non-integrating vectors (**Table 1**).

Table 1. **Characteristics of common viral vectors used in gene therapy.** Adapted from Cabanes *et al.*[31] and Bouard *et al.*[32]

		Adenovirus (Ad)	Adeno-associated virus (AAV)	Retrovirus/Lentivirus
Particle Characteristics	Family	Adenoviridae	Parvoviridae	Retroviridae
	Genome	dsDNA	ssDNA	ssRNA (+)
	Capsid	Icosahedral	Icosahedral	Icosahedral
	Coat	Naked	Naked	Enveloped
	Virion Polymerase	-	-	+
	Virion Diameter	70-90nm	18-26nm	80-130nm
	Genome Size	38-39 kb	5 kb	3-9 kb
Gene Therapy properties	Infection or Tropism	Dividing and non -dividing cells	Dividing and non -dividing cells	Dividing and non -dividing cells
	Host Genome Interaction	Non-integrating	Non-integrating [(b)]	Integrating
	Immunogenicity	High (1st and 2nd generation)	Low	Low
	Transgene expression	Transient	Potentially long lasting	Long -lasting
	Packaging capacity	30 kb [(a)]	4.4 kb	8 kb

[(a)]For gutless adenoviral vectors the packaging capacity is expanded to 30kb. [(b)]AAVs are predominantly maintained episomally although integration at low frequency has been described[32].

Integrating vectors, such as those based on retroviruses and lentiviruses, incorporate transgenes randomly or semi-randomly into the host genome. These vectors were used in initial gene therapy clinical trials and are still used to target dividing cells preserving the therapeutic transgene upon cell division. Although integration ensures stable transgene expression in the target cells and its progeny, it also introduces risk of insertional mutagenesis[27]. The development of engineered self-inactivating (SIN) retroviral vectors, mitigate the risk of integration enhancing their safety profile, particularly in *ex-vivo* applications[30]. This explains their abundant use in gene therapy clinical trials being the second vector of choice (**Figure 2A**).

In contrast, non-integrating vectors, such as those based on adenovirus (AdV) and AAV, are the preferred methods for targeting slow or non-dividing cells, like muscle cells or neurons[32]. Among these, AdV-based vectors have been the most extensively utilized, with ongoing engineering efforts aimed at eliminating early adenoviral genes to reduce immunogenicity enabling the introduction of larger transgene within the vector, with gutless vectors capable of including 30 kb of therapeutic genetic material (**Table 1**).

Over time, AAVs have become the preferred choice for viral gene therapy applications due to their favorable characteristics compared to other viral vectors (**Figure 2B, Table 1**). Despite their limited

capacity for carrying genetic material compared to adenoviral vectors (4.4 kb), vector genomes remain episomal after targeting cells, resulting in transient gene transfer in dividing cells but potentially long-term in non-dividing tissues, with a low degree of insertional mutagenesis, and with a low immune response compared to adenovirus-based vectors, contributing to their excellent safety profile[32].

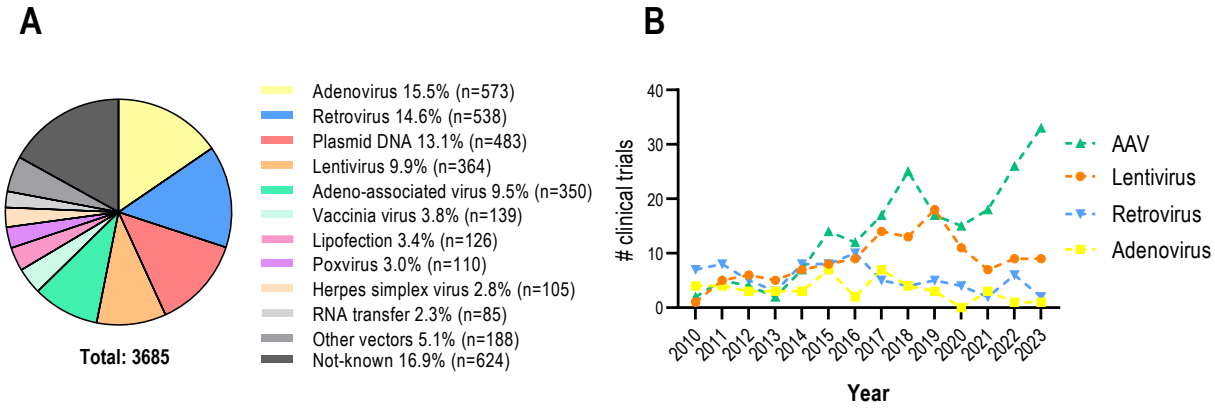


Figure 2. **Trends on using different vectors in gene therapy clinical trials. A.** Prevalence of gene transfer vectors used in clinical settings. Adenoviral, retroviral and naked plasmid/DNA have been the most commonly used gene transfer vectors. Data from Ginn *et al*.[26]. **B.** Time trend of gene therapy clinical trials using lentivirus, adenovirus, AAVs and retrovirus, adapted from Arabi *et al*.[2]. AAVs are the preferred choice nowadays.

2. Adeno-associated virus (AAV)

AAVs were discovered by Atchinson *et al*. in 1965 as a contaminant virus in wild-type adenovirus preparations[33]. Their replication was found to be dependent on the presence of adenoviruses, hence named adeno-associated viruses[33,34]. Due to their replication-defective nature, they were classified under the *Dependovirus* genus within the *Parvoviridae* family.

A major milestone in the use of AAVs as a gene delivery vector was the successful cloning of the AAV genome into a plasmid by Samulski *et al*. in 1982, followed by Laughlin *et al*. a year after[35,36]. This facilitated the development of AAVs as a new type of genetically modified viral particle, capable of carrying therapeutic DNA. Following this, the ability of AAVs as a vector to transduce mammalian cells was demonstrated, being the start of using AAVs as a toolbox for gene delivery[37,38].

Today, recombinant AAVs (rAAVs) are considered the preferred choice in the gene therapy field over other vectors for their desirable characteristics[3]. These include their non-pathogenic nature[39], low immunogenicity, the ability to transduce both dividing and non-dividing cells, and notably, their broad tropism. This tropism is achieved through the diversity of AAV serotypes, which are determined by capsid proteins. Each serotype can infect different cell types, thereby expanding the potential applications of rAAVs in gene therapy. The capsid modifications, either through natural

variation or engineered modifications, allow rAAVs to target specific tissues or evade immune detection, enhancing their efficacy and safety as gene delivery systems.

Furthermore, a key attribute of rAAVs is their potential to induce robust and long-lasting transgene expression by remaining in episomal form within the nucleus of target cells, contributing to stable transgene expression without integrating into the host genome[40]. Although wild-type AAVs are known to integrate into a specific site on chromosome 19 (AAVS1) in humans, the likelihood of rAAV integration is considered to be extremely low, due to the removal of viral genes necessary for this process. To date, no evidence directly links rAAV use to tumorigenesis in clinical settings[41].

In the following sections, the genes and their associated biological functions within both wt AAV and AdV genomes are reviewed. Unravelling these interactions is crucial for advancing the development of rAAV production systems.

2.1. Adeno-associated virus genome and structure

AAVs are formed by a single-stranded DNA (ssDNA) genome of 4.7 kb, enclosed by a 26 nm icosahedral capsid (**Figure 3A**). The AAV genome is organized into three overlapping open reading frames (ORFs), encoding gene sequences for the expression of structural and regulatory viral proteins. These are flanked by 145 nucleotide (nt) long T-shaped hairpin-like structures at both ends, named inverted terminal repeats (ITR)[42,43]. The ITRs play a crucial role in AAV life cycle, providing self-priming activity for DNA replication and packaging signal to encapsidate the ssDNA into the capsid[35,44].

2.1.1 Adeno-associated virus genome

2.1.1.1. AAV replication gene: large and small Rep proteins

The *rep* gene, encoded by the left ORF of the AAV, mediates the expression of four nonstructural proteins that are multifunctional and play an essential role in viral life cycle, including genome replication and AAV encapsidation (**Figure 3B**)[42,45].

The large Rep proteins (Rep78 and Rep68) are transcribed from the p5 promoter[46,47]. Both proteins share a common N-terminal region but differ in the C-terminal due to alternative splicing of their common transcript. Both Rep78 and Rep68 proteins exhibit DNA binding, helicase and

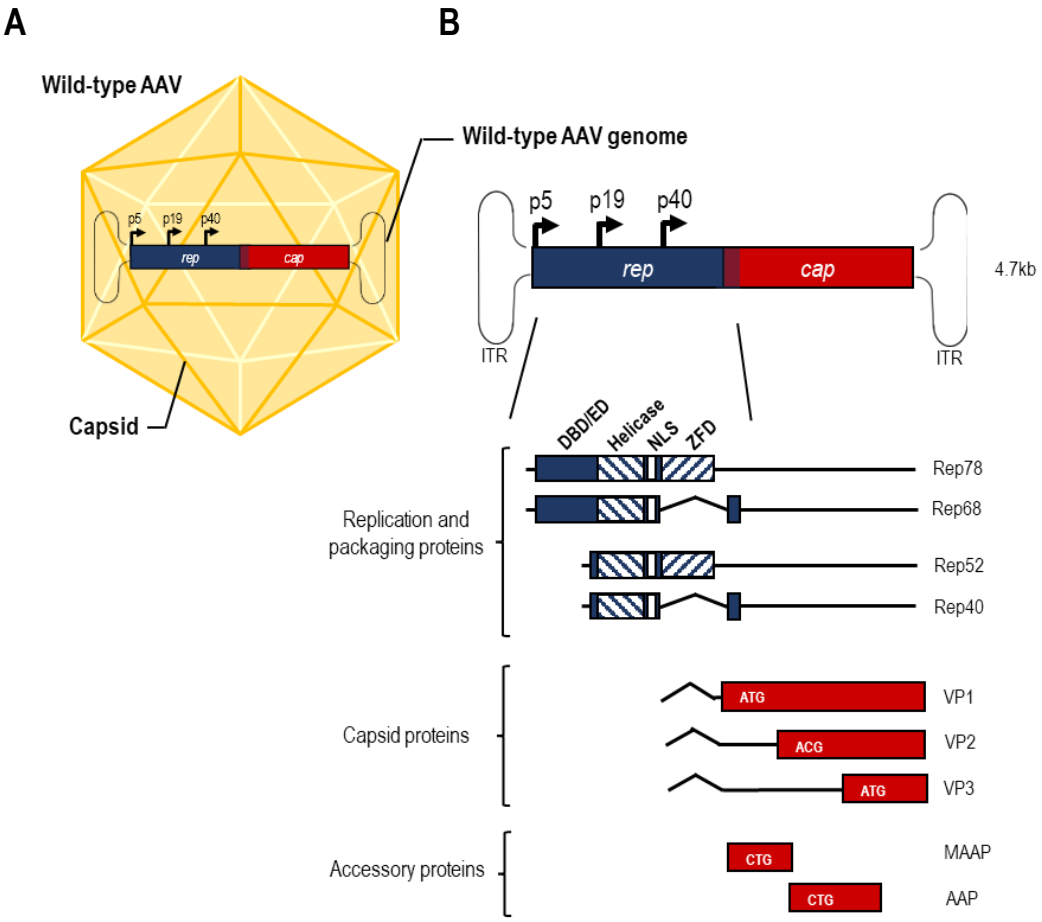


Figure 3. **Wild-type AAV genome structure. A.** The wild-type AAV capsid is of 26nm and contains a single-stranded DNA (ssDNA) genome of ~4.7 kb of flanked at both ends by inverted terminal repeats (ITRs). **B.** The AAV genome encodes four non-structural proteins (Rep78, Rep68, Rep52 and Rep40) and structural proteins forming the capsid (VP1, VP2 VP3). Additionally, MAAP and AAP produced by alternative ORF of the p40 promoter. DBD, DNA-binding domain; ED, endonuclease domain; NLS, nuclear localization site; Wt, wild-type; ZFD, zinc finger domain.

endonuclease activities[48,49]. Their sequence-specific DNA binding domain drives AAV DNA replication by binding to the ITR at the Rep binding site (RBS), which consists of several direct repeats of 5'-GCTC-3' motif[50,51]. Then, the terminal resolution site (trs) is nicked though the endonuclease activity, forming a lineal strand that serves as the origin of replication[48,52,53].

Similarly, small Rep proteins (Rep52 and Rep40) are expressed from the downstream p19 promoter encoded within Rep ORF. Although they lack the DNA-binding domain, are essential for the efficient packaging of the AAV genome into capsid structures via the ATP-dependent helicase activity[54,55] and for ssDNA accumulation during replication[56].

Both Rep68 and Rep40 proteins are produced by alternative RNA splicing of Rep78 and Rep52 transcripts, respectively. They lack the 91 C-terminal amino acids corresponding to the zinc finger domain (ZFD) and have an additional seven amino acids encoded by exon 2[57].

Notably, the large Rep proteins have a strong antiproliferative effect, less pronounced for the small Rep proteins[57]. Rep78 has been shown to induce cell death through a caspase-3 mediated apoptosis, independently of p53 in a cell cycle-dependent process[58]. Particularly, Rep78 and Rep68 block cell cycle progression differently due to the differential ZFD: while Rep78 induces S-phase arrest with the accumulation of hyperphosphorylated retinoblastoma protein (pRb), Rep68 induces arrest at both G1 and G2 phases[57,59]. This pro-apoptotic activity of Rep proteins poses a significant challenge in rAAV production systems, which can severely impact the viability of rAAV-producer cells. In this work, the Rep protein expression is aimed to be inducibly controlled, aiming to decrease its toxic effects towards the development of stable rAAV production strategies.

2.1.1.2. AAV capsid gene: VP proteins

The transcription of the *cap* gene is controlled by the p40 promoter encoded, encoding the capsid proteins VP1, VP2 and VP3 (**Figure 3B**). The VP1 is produced from the full cap transcript translation, while VP2 and VP3 are the product from alternative splicing and their translation is initiated from an alternative start codon (ACG and ATG, respectively)[60,61].

2.1.2 Adeno-associated virus capsid structure and serotypes

The AAV capsid is characterized by a highly symmetric geometry, formed by an icosahedral capsid structure consisting of 60 VP units, assembled in a stoichiometric ratio of 1:1:10 of VP1, VP2 and VP3 respectively[61]. The unique N-terminal region of VP1 is essential for virus infectivity, owing to its phospholipase A2 (PLA2) activity and the presence of a nuclear localization signal (NLS)[62]. The N-terminus of VP2, on the other hand, has been found to be nonessential for viral infectivity and it has the capacity to tolerate peptide insertions, facilitating cell-specific targeting of AAV[63]. Particles consisting of VP3 alone can form capsid structures, although not infectious[62].

The discovery of various AAV serotypes, differentiated by variations in capsid sequences, has expanded the applications of AAV in gene therapy[64]. To date, 13 AAV serotypes and over 100 AAV variants have been identified and classified from a variety of species including mammals, birds and reptiles[65]. Each serotype is associated with its ability to bind to a specific cellular receptor, tissue tropism and antigenicity based on its capsid structure.

The amino acid sequence identity in the VP1 region among these serotypes ranges from 57% to 99%[66]. However, the common VP region maintains a highly conserved structure across serotypes, characterized by a β-barrel motif core forming the capsid interior[67]. This core is connected by loops on the capsid surface, which exhibit variable conformation among serotypes, attributable to differences in the hypervariable region of the VP2 protein conferring serotype-specific features[68,69].

The common structural topology across the AAV serotypes includes depressions at the twofold symmetry axis, three protrusions at the threefold axis surrounding the depressions, and pores or channels at the fivefold axes, that allow exchange between the capsid interior and the outside[66,67,70].

2.1.3 Adeno-associated virus receptors

The receptor binding affinity of the AAV capsid is a key characteristic influencing tissue tropism and distribution. Generally, AAV serotypes bind to cell surface glycans or glycoconjugates, essential for viral transduction[71]. Early studies by Samulski *et al*. demonstrated that AAV2 binds to heparan sulfate proteoglycan (HSPG) (**Table 2**)[72]. The widespread expression of HSPGs explains the broad tropism observed for AAV2. Besides primary receptors, AAVs bind to secondary receptors and co-receptors, enhancing their concentration which facilitates viral entry[73]. For AAV2, a number of co-receptors have been identified, including human fibroblast growth factor receptor 1 (FGFR1), αVβ5/α5β1 integrin, hepatocyte growth factor receptor (HGFR), and laminin receptor (LamR)[73,74]. To date, different receptors and co-receptors for most serotypes have been identified (**Table 2)**.

In addition, Pillay *et al*. in 2016 identified an essential AAV2 receptor, known as AAVR[75]. Subsequent findings revealed that AAVR is essential for infection of several AAV serotypes, including AAV1, 2, 3, 5, 6, 8 and 9[75]. Today, it is known that the different serotypes exhibit varying binding affinities to AAVR and it is not required for AAV4 infection[76].

2.1.4 Prevalence of neutralizing antibodies against AAVs

The exposure of humans to AAV is highly prevalent. Although no pathology is associated with AAV infection in humans, the AAV capsid can trigger an immune response leading to pre-existing humoral immunity, including neutralizing antibodies against wt AAVs. Variations in the seroprevalence of anti-AAV antibodies have been demonstrated across different geographic regions and serotypes[77]. In a study conducted in France, the seroprevalence of neutralizing factors against specific AAV serotypes showed that anti-AAV2 antibodies exhibited the highest prevalence,

reaching up to 60% in healthy population (**Table 2**)[78]. Recent studies reported even higher prevalence rates for antibodies against AAV8 and AAV2, ranging from 40% for anti-AAV8 to 74% for anti-AAV2[79]. Additionally, the high degree of homology between AAV capsids facilitates cross-reactivity among neutralizing antibodies, complicating the strategy of switching serotypes to evade immune detection[79,80]. The presence of anti-AAV neutralizing antibodies is undesirable for rAAV-based gene therapy administration as this antibodies can inhibit vector transduction and its clinical effect[39].

Table 2. Cellular receptor and co-receptors, tissue tropism and antigenicity of different AAV serotypes. Compilation of receptors and co-receptors for AAV entry, tissue tropism and prevalence of neutralizing antibodies to AAV capsid in humans. Adapted from Tejash *et al*.,[81] Cabanes *et al*.,[31] and Verdera *et al*.[73].

Serotypes	Receptors and Co-Receptors	Tissue Tropism	Neutralizing factors prevalence[(a)]
AAV1	α2-3/α2-6 N-linked SA	skeletal muscle, lung, CNS, retina, pancreas,	50.5% (n=152)
AAV2	HSPG, FGFR1, HGFR, Integrin αVβ5/α5β1, LamR, CD9	Smooth muscle, skeletal muscle, CNS, liver, kidney	59% (n=89)
AAV3	HSPG, FGFR1, LamR, HGFR	Hepatocarcinoma, skeletal muscle, inner ear	ND
AAV4	α2-3 O-linked SA[(b)]	CNS, retina	ND
AAV5	α2-3 N-linked SA, PDGFR	Skeletal muscle, CNS, lung, retina, liver	3.2% (n=49)
AAV6	HSPG, EGFR, α2-3/α2-6 N-linked SA	Skeletal muscle, heart, lung, bone marrow	37% (n=56)
AAV7	ND	Skeletal muscle, retina, CNS	ND
AAV8	LamR	CNS, liver, muscle, heart, pancreas	19% (n=50)
AAV9	N-linked galactose, LamR, Putative integrin	CNS, heart, lung, liver, muscle	33.5% (n=62)

[(a)]Relative to the number of tested samples from Monteilhet *et al*.[78]. [(b)]AAVR independent. AAVR, Adeno-associated virus receptor; CNS, Central nervous system; EGFR, epidermal growth factor receptor; FGFR1, fibroblast growth factor receptor-1; GAL, galactose (terminal N-linked); HGFR, hepatocyte growth factor receptor; HSPG, heparin sulfate proteoglycans; LamR, Laminin receptor; ND, not determined; PDGFR, platelet-derived growth factor receptor; SA, sialic acids.

2.1.5 Adeno-associated virus accessory proteins

Additional nonstructural accessory proteins, expanding the coding capacity of the AAV genome have been identified through alternate ORF within the *cap* gene (**Figure 3B**).

In 2010, Sonntag *et al*. described the assembly-activating protein (AAP), expressed from a +1 frameshifted ORF within the *cap* gene, under the p40 promoter control[82]. AAP from AAV2 (AAP2), has been the most studied to date. In cultured cells, AAP2 and AAV2 VP3 proteins interact through

hydrophobic domains, enabling AAP function as both a scaffolding protein in the capsid assembly reaction and as a transporter that targets VP proteins to the nucleolus for assembly[83–85]. Additionally, the presence of AAP has been shown to increase capsid protein stability and VP-VP interactions[86]. Interestingly, studies have suggested that AAP may play a role in providing an evolutionary advantage by favoring capsid evolution[86].

Of note, significant serotype-dependent heterogeneity has been observed in the AAV capsid assembly process, particularly regarding the AAP requirement and the sub-localization within the nucleus. For example, while AAV2 shows nucleolar enrichment, AAV5, 8, and 9 are excluded from this nuclear compartment[87]. Furthermore, AAV4, 5, and 11 capsid assembly does not depend on AAP[85].

A second accessory protein encoded in a different reading frame from the VP1 ORF was identified by Ogden *et al*.[88]. This ORF encodes a protein reported to be associated with cell membranes, such as plasma membrane, perinuclear membranous structures, and nuclear envelope, hence named "membrane-associated accessory protein" (MAAP). Recent findings have supported a functional role for MAAP in mediating AAV egress in rAAV studies, reporting that reducing MAAP8 expression, results in a delay in extracellular secretion of viral particles, leading to increased cellular retention[89]. This process has been demonstrated to be mediated through the association of MAAP with the outer surface of extracellular vesicles (EVs), dependent on a cationic and hydrophobic domain in MAAP crucial for membrane anchoring. Therefore, the association of MAAP with EVs has been shown to alter AAV egress and kinetics, promoting the extracellular secretion of particles[89].

2.2. Adenoviral helper functions for adeno-associated virus replication

AAVs can infect host cells both in the absence and presence of helper factors. However, as members of the *Dependovirus* genus, they require helper virus co-infection for replication and propagation. Adenovirus serotype 5 (AdV5) is the most extensively characterized helper virus, through Herpes Simplex Virus type 1 (HSV1) and baculovirus have also been shown to provide complete helper functions to AAVs in cell culture[90–92]. Several other helper viruses have been identified though they are less well studied. For example, the Human Papillomavirus type 16 (HPV16) has been shown to promote AAV replication via *E1, E6* and *E7* genes[93]. Similarly, the human bocavirus 1 (HBoV1), member of the *Parvoviridae* family, has been shown to provide helper functions[94]. Other viruses, such as vaccinia virus, have been reported to provide subhelper function, not fully supporting but promoting AAV replication[95]. Additionally, in the absence of other

viruses but under genotoxic stress from physical or chemical agents, such as topoisomerase inhibitors or UV irradiation, the AAV replication can also be induced, although the mechanisms behind this are still unknown[95,96].

The molecular mechanisms providing AdV helper functions on AAV life cycle is provided hereafter. Noteworthy, the AdV does not supply proteins directly involved in AAV DNA replication. Instead, AAV relies on cellular replication proteins and the adenovirus provides a range of support functions. These include activating AAV promoters that would otherwise be repressed, modifying the host cell environment and interacting with cellular factors to enhance the cell's permissiveness to AAV replication[97]. The set of adenoviral genes essential for AAV2 replication and propagation consists on *E1A, E1B, VA RNAs, E2A* and *E4orf6*[90,98–101] (**Figure 4**, **Table 3**).

The characterization of this gene's functions is important towards designing rAAV production methods based on the AdV helper complementation.

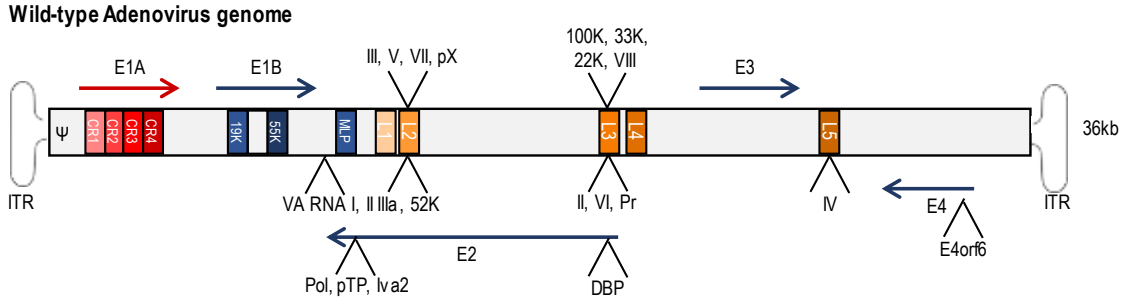


Figure 4. **Schematic representation of the wild-type Adenovirus type 5 (AdV) genome**. The 36 kb genome consists of four early transcription genes (*E1, E2, E3* and *E4*), five late expression genes (*L1-L5*) driven by the major late promoter (MLP), cis-packaging elements (Ψ) and two inverted terminal repeat sequences (ITR). The *E1A* gene contains four conserved domains (CR1-4), each of which interacts with cellular proteins. The *E1B* gene encodes two proteins (19K and 55K with 19 and 55 kDa respectively). The *E2* gene encodes the DNA-binding protein (DBP), the terminal protein (pTP), Iva2 and Pol. The *E4* gene encodes up to 7 different ORF. The virus-associated RNA (VA RNA I and II) are noncoding RNAs. Adapted from Bulcha *et al.*[102].

Upon adenovirus co-infection, the E1A proteins encoded by the *E1A* gene are the first to be transcribed, acting as a cue to initiate AAV replication. First, E1A upregulates transcription from the AAV p5 promoter, which controls the expression of large Rep protein[47,103,104]. Additionally, E1A transactivates early adenovirus promoters, including *E1B, the VA RNAs, E2A, and E4orf6*[99,101,105–109].

Beyond its role in viral replication, E1A manipulates the host cell environment by modulating host cell proteins to favor AAV replication, specifically by promoting entry into the S-phase of the cell cycle[110]. Contrarily to this oncogenic function, E1A can also act by triggering apoptosis both by a

p-53 dependent and p-53 independent mechanisms[111]. In the context of adenovirus-infected cells, the *E1B* gene products (19K and 55K proteins) exhibit anti-apoptotic functions[112,113], presumably to counteract E1A pro-apoptotic properties and ensure productive infection and cellular transformation[114,115].

In AAV replication settings, the E1B 55K protein has been reported to be essential, mainly in complex with E4orf6, forming a helper complex factor with different functions described in **Table 3**. This complex is involved in promoting AAV second-strand DNA synthesis[116], exporting AAV mRNA from the nucleus to the cytoplasm[117], and promoting the degradation of the MRN complex thereby promoting AAV replication in host cells[118–120].

Additionally, the adenovirus *E2A* gene, also transactivated by E1A, encodes various proteins, among them the single stranded DNA-Binding Protein (DBP). The DBP can also induce p5 AAV promoter, similar to E1A protein[121]. DBP plays a pivotal role in increasing intracellular levels of AAV genome, stimulates AAV DNA replication and capsid protein production[61,122–125], and it is involved in other functions such as Rep mRNA splicing and export from the nucleus to the cytoplasm[121,123].

Notably, both E2A and E4 products have shown toxicity to host cells. Despite the toxicity of the DBP to human cells, constitutive expression of E2A has been achieved in clones derived from HEK293 cells[126,127]. Regarding E4 products, particularly E4orf4 expression has been documented to induce cellular toxicity through induction of p-53 independent apoptosis[128], while the E4orf6 in complex with the E1B 55K protein are reported to mediate the degradation of p-53 together counteracting the toxic effect[129]. Of note, constitutive expression in human cell lines of E4 products has not been achieved to date, indicating its remarkable cytotoxic potential.

Finally, virus-associated RNA (VA RNAs) are noncoding RNAs involved in enhancing viral protein translation by circumventing the host cell's antiviral response in a complex range of viral-host dynamics detailed in **Table 3**.

In the context of rAAV vector production, the requirement of helper adenovirus proteins is similar as in the wt AAV virus replication. However, the indispensability of specific E1B 19K or 55K proteins in rAAV production has not been fully elucidated to date, with conflicting evidence reported. In a study by Matsushita *et al*. characterizing essential functions of the *E1* gene in rAAV production, reported that while E1B 19K increased rAAV production, E1B 55K did not play a significant role, indicating no direct involvement of E1B 55K in the rAAV production process[101]. In addition, anti-apoptotic factors homologues to E1B 19K were able to fully complement E1B 19K mutants for rAAV production[101]. This is in conflict with observations by Samulski *et al.* in which AAV helper function

was attributed to the E1B 55K protein[130]. This underscores the complexity in translating adenoviral helper functions from natural AAV production to rAAV vector production settings as the later study was performed in the context of AAV and AdV co-infection, while the former was based on rAAV production. Therefore, this indicates that caution must be taken when correlating all AdV helper functions studies towards rAAV production requirements.

Table 3. Description of essential adenoviral helper proteins and non-coding RNA functions essential for AAV replication. Adapted from Meier *et al*.[90].

AdV helper element	Helper function for AAV replication	References
E1A	AAV p5 promoter activation	47,103,104
	Relieve p5 repression by binding its repression factor YY1	131
	AdV early promoter activation	99,101,105–109
	Driving cells to S-phase	110
E1B 19K[(a)]	Rescue of latent AAV and enhancement of replication	101
E1B 55K - E4-orf6 complex	Prevents E1A-mediated p53 stabilization, targeting p53 for degradation	112–115 129
	Inhibits MRN complex	118–120
	Promotes second strand AAV synthesis	116 116
	Favours Rep mRNA splicing and export from the nucleus to cytoplasm	121,123
DNA-Binding Protein (DBP)	AAV p5 promoter regulation	121
	ssDNA binding protein, viral DNA replication	122–125
	Increase capsid protein production	61
	Improve mRNA rep processing (protein splicing)	121,123
VA RNA I and II	Inhibiting eIF-2 protein kinase, promotes viral protein translation	132
	Inhibition of host PKR, facilitating AAV protein synthesis	133,134

[(a)] In AAV and AdV coinfection studies, only E1B 55K was essential[101]. However, in the context of rAAV production E1B 19K and not E1B 55K is required[130].

2.3. Adeno-associated virus life cycle

a) Cellular attachment and endocytosis

The AAV life cycle begins with cell recognition and attachment of the capsid surface to negatively charged glycans. As described, different serotypes exhibit affinity for different cell receptors and co-receptors. The AAV2 life cycle, which is the best studied serotype, is described hereafter (**Figure 5**, step 1 and 2).

Following AAV2 attachment to HSPG, AAVR, and various co-receptors, endocytosis is initiated by the invagination of the plasma membrane. AAVR not only assists in virus attachment, but also in its internalization process from the plasma membrane to the Golgi network[135]. For AAV2, this is a rapid and efficient process occurring within about 30-60 min[136]. Currently, some discrepancy exists regarding the precise AAV endocytic pathway, which is expected to vary between serotypes. Initial studies suggested a clathrin-mediated endocytosis pathway dependent on dynamin for AAV2[137]. However, subsequent research has indicated that AAV can use alternative entry routes, independent on clathrin such as the clathrin-independent carriers/glycophosphatidylinositol-enriched early endosomal compartments (CLIC/GEEC) pathway[138].

b) Endosome trafficking, release, and nuclear entry

Following the internalization, AAV traffics along the endo-lysosomal pathway via the trans Golgi network and towards the microtubule organizing center (MTOC)[139–141] (**Figure 5**, step 3 and 4). Then, the endosome vesicle undergoes acidification that triggers capsid conformational changes exposing the N-terminal domain of the VP1, and the PLA2 activity disrupts the lipid membrane of the endosome and mediates viral particle escape to the cytosol[62,142,143]. After escape, AAV particles travels to the nucleus periphery and enters through the nuclear pore complex (NPC) as intact particles[144]. This mechanism is known to be mediated by the canonical nuclear import pathway, in which the NLS at the VP1 and VP2 N-terminal region bins to importin-β alone or in complex with other karyopherins[145].

c) Genome uncoating and expression

Upon reaching the nucleus, the viral particles are disassembled and the viral genome is released in a process not fully known[144] (**Figure 5**, step 5). The current studies proposes a model where AAV2 particles accumulate in the nucleolus upon nuclear entry, and the uncoating of AAV2 is induced by the structural reorganization of the nucleolus in a cell cycle-dependent manner[146].

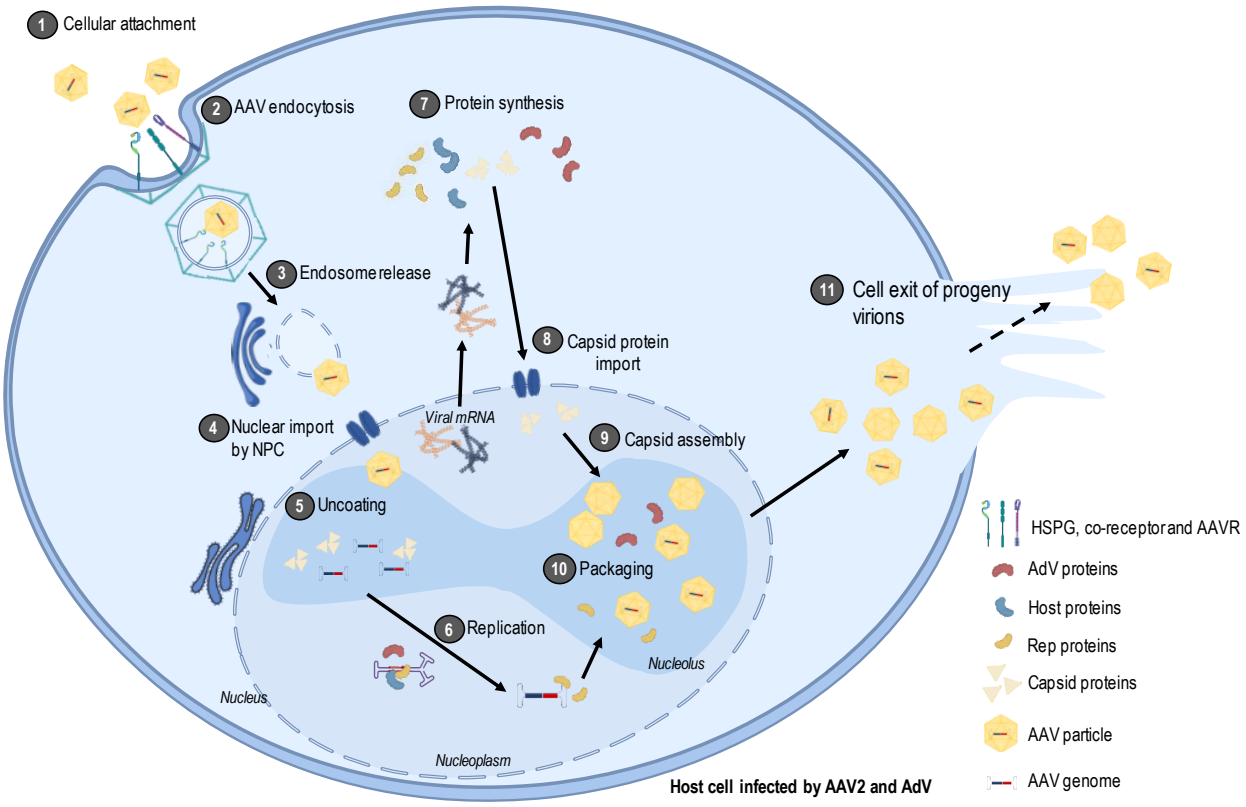


Figure 5. **Schematic representation of AAV2 replication by AdV co-infection.** 1. Cellular attachment of AAV2 to AAVR, HSPG and other co-receptors. 2. AAV endocytosis by clathrin pathways. 3. Endosome release by pH shift triggering capsid conformation change. 4. Nuclear import by the nuclear pore complex (NPC). 5. AAV uncoating. 6. AAV replication detailed in **Figure 6**. 7. Protein synthesis. 8. Capsid protein import. 9. Capsid assembly. 10. Packaging of ssDNA to the preformed capsids. 11. AAV release.

d) Lytic cycle: replication and packaging

Upon AAV uncoating, whether the AAV life cycle process proceeds into a lytic or latent phase is dependent on the presence of helper proteins, such as those provided via adenovirus co-infection. In the presence of adenovirus helper functions, AAV DNA replication proceeds (**Figure 5**, step 6 to 10). This process requires AdV helper factors and AAV replication proteins, as described in previous sections, along with host cell replication machinery. Essential host factors for AAV DNA replication have been identified, including DNA polymerase δ (pol δ), minichromosome maintenance (MCM) helicases, replication factor C and A (RFC, RFA) and proliferating cell nuclear antigen (PCNA)[50,90,145].

The replication process is driven by the self-annealing property of the ITRs, which provides a free 3'OH end acting as an origin for viral replication (**Figure 6A**). This initiates the "rolling hairpin mechanism" of replication coupled with strand displacement mechanism of resolution, which unfolds in several steps schematically represented in **Figure 6B** and described below[147]:

1. Cellular pol δ identifies the free 3'OH end of the ITR as a primer, initiating unidirectional DNA elongation (green). This results in dsDNA that is capable of transcribing the AAV *rep* and *cap* genes, which then assist in the AAV replication process and form AAV capsids, respectively.
2. Large Rep proteins bind to the RBE motif within the ITRs and introduce a nick at the trs. The dsDNA is converted to ssDNA and large Rep proteins attach to the 5' end along with host pol δ and AdV helper factors.
3. The combined action of these proteins unwinds the DNA, providing a new 3'OH for extending the ssDNA. This extension replicates the ITR on the daughter strand (orange), generating two linear strands as a result (replication intermediates).
4. Subsequent renaturation of the ITR hairpin structures on both strands occurs, facilitating the reformation of secondary structures.
5. The free 3'OH end guides single strand displacement synthesis, resulting both in a dsDNA and a full ssDNA genome product. The dsDNA may be used as a template to re-initiate replication process.
6. Finally, the ssDNA genome is packaged into preformed capsids forming new virion particles.

As described, the AAV capsid assembly process is serotype-specific, with capsid protein oligomerization occurring in the cytoplasm and assembly in the nucleolus[148]. Once assembled, the newly synthesized ssDNA is packaged into the preformed capsids by the helicase/ATPase domain of the small Rep proteins in a 3' to 5' direction, translocating through the fivefold channels[55]. DNA packaging likely occurs separately, in the nucleoplasm[84].

Interestingly, AAV replication, in the context of adenovirus co-infection, has been identified to occur in viral replication centers localized in the nucleoplasm, where AdV and cellular proteins are available for AAV replication[124,146].

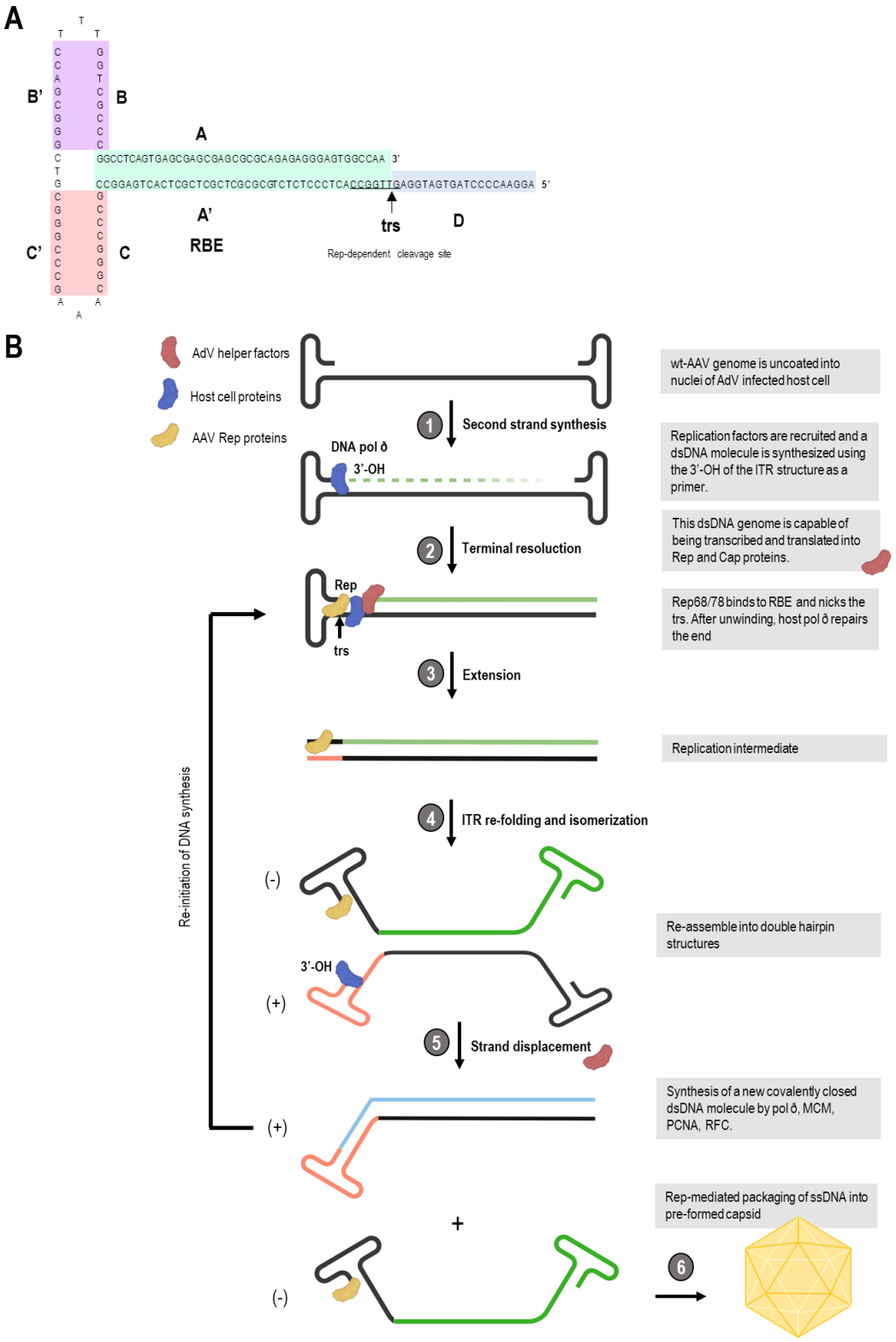


Figure 6. AAV DNA replication. A. Basic features of AAV ITR consisting of two small palindromes with complementary sequences (B and C) flanked by a large palindrome (A) and 20 additional 20-base sequence (D) repeated at both ends of the AAV genome. **B.** Schematic representation of AAV ssDNA replication by the strand displacement mechanism. New generated AAVs DNA segments are colored in green, orange and blue. Displaced AAV genomes are encapsidated into preformed capsids. AdV helper proteins are represented in red. Host cell proteins are represented in blue. AAV Rep proteins are represented in yellow. ITR, inverted terminal repeat. trs, Terminal Resolution site. Adapted from Bennett *et al.*[149] and Cabanes *et al.*[31].

e) Cell secretion of AAVs

Once the lytic cycle is completed, progeny virions must exit the cells to complete the infectious cycle[150]. Typically, enveloped viruses generally exit via exocytosis, while non-enveloped viruses are traditionally assumed to egress by cell lysis[150].

For AAVs, the exit mechanisms are not yet fully deciphered (**Figure 5**, step 11). Unlike other parvoviruses that undergo a lytic cycle, AAVs do not induce significant cytopathic effects (CPE). Instead, AAV release is thought to be primarily driven by the overexpression of adenoviral or herpesvirus helper proteins at the end of an infection[89,90,150].

However, Lock *et al.* and Vandenberghe *et al*. in 2010 first characterized that rAAVs are capable of being secreted into the cell culture media before cell lysis, with variable efficiencies depending on the serotype[151,152]. Further, they demonstrated a correlation between the rAAV concentration in the medium and its binding affinity to HSPG; AAV2, with high HSPG affinity, was mainly found in the cell pellet, whereas AAV8 was predominantly in the medium[72].

Additionally, several studies have identified that a fraction of rAAV isolated from the extracellular medium are associated with exosomes, termed exo-rAAV[153,154]. These exo-rAAVs exhibit enhanced transduction efficiency compared to conventional rAAV vectors[155]. Notably, the MAAP protein, reported as a viral egress factor, promoted rAAV exit from the cells by interacting with the surface of extracellular vesicles[89]. The same study reveals that when MAAP8 expression is reduced, it causes a delay in the extracellular secretion of AAV particles, leading to increased cellular retention of the particles[89].

This evidence underscores that rAAV secretion pathways such as the related to the vesicular transport needs to be further elucidated[156]. Understanding these mechanisms holds significant implications for maximizing the rAAV at the extracellular fraction, enabling the harvest of viral vectors at the extracellular media, therefore, with lower quantity of cell-derived products compared to the cell lysate**.**

f) Latent cycle: episomal persistence and specific integration

In the absence of AdV co-infection, AAV establish a latent infection either through long-term episomal persistence or by preferential integration of the viral genome into specific sites in the human genome. The mechanism of AAV2 latency is mediated by two approaches. First, large Rep proteins mediate autorepression of AAV2 genome expression by binding the RBE in the p5 promoter in conjunction with cellular factors YY1 and MLTF, thus suppressing its transcription[131].

In addition, Rep protein binds to the KAP1 cellular enzyme, leading to the recruitment of chromatin-modifying enzymes that promote heterochromatin formation, to evade the immune system[157]. Using these mechanisms, the AAV2 genome remains silenced in host cells until AdV infection occurs. Upon AdV infection latent AAV can be rescued with the E1A protein, which interaction with p300, enables interaction with YY1 and relieves the p5 repression[131,158].

The mechanism of AAV long-term persistence is mainly through circular extrachromosomal elements[159]. After the synthesis of dsDNA, the genome can undergo circularization via intra-molecular or inter-molecular recombination at the ITRs[160,161]. These can form head-to-tail concatemers if the 3‘ hairpin is not resolved by cleavage at the closed end at the trs, which would require Rep proteins[162].

During latency, the AAV genome can also persist by integration, mostly detected in dividing tissues. The integration mechanism is mediated by AAV Rep proteins which catalyze site-specific integration of the AAV genome into chromosome 19 at the AAVS1 locus[163,164]. This locus contains RBS and trs-like motifs, which are recognized by large Rep proteins that mediate site-specific integration in a mechanism similar to the replication process[165–167]. Several studies have reported AAV2 integration near the AAVS1 site and also other identified hotspots, mostly in RBE repeats, depending on specific chromatin accessibility and cell type[168].

Noteworthy, the rAAV vector, lacking the AAV genome, is replication-deficient and can achieve long-term expression through mechanisms like wt AAV. Once rAAV reaches the interior of target cells, the ITRs drive the intermolecular recombination of the dsDNA to form dimers and concatemers presumably facilitated by host cell enzymes[160]. These episomal DNA structures remain in the nucleus as long as the cell does not divide and mediate long-term transgene expression presenting a low risk of genotoxicity[169,170]. However, studies in non-human primates have shown that rAAV can integrate into the host genome at a low frequency[171]. Although vector integration events are rare and usually not linked with toxicity, some associations with oncogenic events in animal models have been shown[172,173]. Consequently, while rAAV circumvents the possibility of Rep-mediated targeted integration, the risk of oncogenesis due to random integration remains a concern, albeit not fully elucidated to date[174].

3. Adeno-associated virus as a gene delivery vector

3.1. Vectorization process

The growing interest in rAAV-based therapies can be attributed to distinct characteristics that make rAAV vectors particularly suited for gene delivery as previously described. This interest is further supported by the straightforward design process of rAAV as a vector (**Figure 7A, B**). Standard elements for rAAV production based on adenoviral helper system are described hereafter, along with modifications aimed at addressing the current limitations of rAAV as a gene therapy vector.

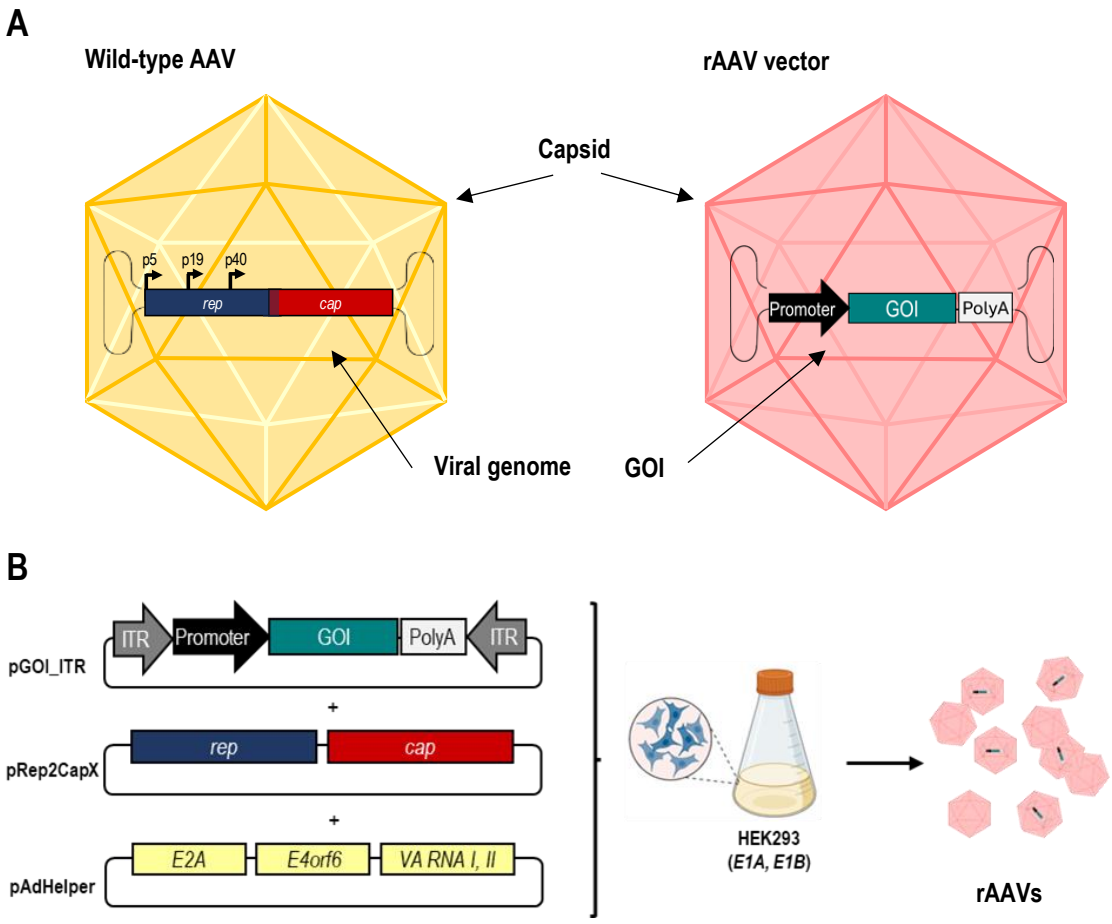


Figure 7. AAV vectorization. A. The viral wt AAV genes (*rep* and *cap*) flanked by ITRs are replaced with the therapeutic gene of interest (GOI) in recombinant AAV (rAAV) genome. The capsid remains intact. **B.** Plasmids used for rAAV production into HEK293 cells by transient transfection method.

3.1.1 AAV expression cassette (pAAV_ITR)

In the vectorization process of rAAV, the wt AAV *rep/cap* genes are removed, leaving a gap of approximately 4.4 kb which is replaced by the therapeutic gene of interest (GOI) (**Figure 7A**). The only wt AAV sequences remaining in the rAAV are the ITRs, typically from AAV2 and serve as the "packaging signal" for the rAAV genomes Therefore, the expression construct consists of the GOI, flanked by a promoter and a termination element, all flanked by ITR sequences (**Figure 7B,** pGOI_ITR). The optimal packaging capacity is between 4.1-4.9 kb, with suboptimal efficiency if the cloned DNA is outside of this range[175]. To overcome the rAAV packaging size limitation, several strategies have been explored. These include genetic engineering approaches to reduce transgene size (e.g., mini dystrophin)[176] and the development of dual-vector systems, such as overlapping or trans-splicing vectors[40]. These systems divide the expression cassette in different vectors and rely on the host cell to reassemble the transgene, although with less efficiency compared to when is delivered in only one vector[40].

Furthermore, the expression of rAAV depends on the transgene design, the presence of regulatory elements, and the choice of the promoter[177]. Several promoters can be used to drive transgene expression, including those derived from other viruses, of mammalian origin, or even those that are cell type-specific to decrease off-target effects. The CMV and CAG promoters are commonly employed in clinical trials across different therapeutic areas for their robust constitutive expression[3]. In contrast, tissue-specific promoters are preferred for blood diseases to minimize off-target activity[3].

3.1.2 Packaging construct (AAV Rep and Cap)

The GOI in the rAAV can be encapsulated into capsids of its own serotype or alternatively, from another serotype through a process called pseudotyping, depending on the desired tropism. For rAAV production, it is common practice to use the *cap* gene from one serotype while utilizing the Rep proteins and ITRs from AAV2. For instance, when employing an AAV8 capsid with AAV2's Rep proteins and ITRs, the resulting vector is termed rAAV2/8. For simplicity, such constructs are named after the capsid serotype, e.g., rAAV8 (**Figure 7B,** pRep2CapX)[99]. Notably, the choice of ITR serotype is closely tied to the Rep sequence used, especially for serotypes like AAV5, where specific Rep-ITR interactions are necessary for efficient viral replication and packaging[178].

As reviewed, one critical consideration in employing different capsid serotypes in the rAAV therapeutic context, is the potential for preexisting antibodies in the host, impacting the vector's

effectiveness. Research is ongoing to identify new variants from natural sources, such as those from the rhesus macaque (e.g., AAVrh.10 and AAVrh.74), which exhibit favorable transduction properties and lower immunity in the human population compared to AAV2[40]. Additionally, novel vectors have been engineered to enhance tissue specificity and modify antigenic sites using capsid genetic engineering techniques[40]. Directed evolution and rational design are the predominant methods for developing these new variants. As of now, although new variants have been tested in pre-clinical studies showing good safety profiles, AAV2 is the most used serotype in clinical settings despite having the highest prevalence of neutralizing antibodies[3].

Therefore, the choice of capsid in rAAV vectors not only affects the vector's tissue specificity and transduction efficiency but also its ability to circumvent immune barriers, making it a critical factor in the successful application of gene therapies. In this context, this work focuses on characterizing and improving the manufacturing processes of rAAV9 and rAAV1, which are selected based on their distinct properties and therapeutic potential.

3.1.3 Adenovirus helper virus construct (pAdHelper)

Finally, adenoviral helper functions are supplied in producer cells. Initially, AdV helper functions were provided by AdV co-infection for rAAV production, presenting safety risks due to the presence of process-related impurities such as contaminating AdV and replication-competent AAV (rcAAV)[179]. rcAAV represent critical contaminants in rAAV products as they are replication competent. During the manufacturing process, rcAAV may originate from recombination events between the rAAV vector and residual AAV replication genes present in the production system. Vector designs that minimize overlap with helper virus sequences may reduce the risk of recombination and it has been subject of study for improving rAAV production systems[180].

A major advancement was the cloning of the essential AdV5 helper elements into a helper plasmid, avoiding the need of AdV helper infection (**Figure 7B,** pAdHelper). This impacted positively the viral yield and safety of the rAAV product[99,106].

Currently, the most common rAAV production method is by triple transfection into HEK293 cells. These cells were originally transformed with the *E1* gene by transfection with sheared AdV5 genomic DNA, resulting in constitutive expression of E1A and E1B proteins[181,182]. Therefore, to provide the full AdV helper gene set, the helper genes are supplied in a plasmid containing *E2A, E4 and VA RNAs* but lack the *E1* gene, as it is already present in HEK293 cell. Additionally, the

described plasmids containing the AAV genes (pRep2CapX) and the GOI (pGOI_ITR) are also supplied.

3.2. Need of cost-efficient rAAV manufacturing systems

In recent years, the field of gene therapy based on rAAV has achieved significant milestones, consolidating it as a platform for treating rare monogenic diseases. Since the approval of the first rAAV-based product, Glybera® in 2012 the rAAV field has constantly expanded and the number of rAAV-based gene therapy clinical trials has increased, with 121 active and recruiting as of September 2023. In addition, five rAAV-based gene therapy products have been approved by the EMA to date (**Figure 8**)[20]. However, the scope of these products in terms of the potential patient population that can be reached, is limited due to their remarkably high prices, constantly breaking price records. For example, Hemgenix®, an AAV5-based treatment for Hemophilia B is the most recently approved product with the potential of replacing the need of lifelong therapies with what is expected to be a single infusion, although at a price of $3.5 million in the United States, making it the world's most expensive therapy.

EMA year of approval	2018	2020	2022	2022	2022
	Luxturna	Zolgensma	Roctavian	Upstaza	Hemgenix
Vector serotype	AAV2	scAAV9	AAV5	AAV2	AAV5
Indication and therapeutic gene	Inherited retinal dystrophy (hRPE65)	Spinal Muscular Atropy, Type I (SMN)	Hemophilia A (Factor VIII)	L-amino acid decarboxylase (AADC) (Dopa-Decarboxylase)	Hemophilia B (Factor IX)
Industry	NOVARTIS Spark THERAPEUTICS	NOVARTIS avexis	BIOMARIN	PTC THERAPEUTICS	uniQure CSL Behring
Dose	$1.5x10^{11}$ vg/eye	$1.1x10^{14}$ vg/kg	$6x10^{13}$ vg/kg	$1.8x10^{11}$ vg/kg	$2x10^{13}$ vg/kg
Price/treatment	$850.000 both eyes	$2.1million	$2.9 million	$3.7 million	$3.5million

Figure 8. Approved rAAV-based gene therapy products by EMA. For each product, the commercial name and year of approval by the European Medicines Agency (EMA) is indicated. Therapeutic indication, the therapeutic gene the sponsoring entity of the drug product, and specifics regarding the required therapeutic dose and the cost per treatment are detailed. Vg, viral genomes.

These elevated costs are partly due to the fact that many gene therapies target rare diseases with limited patient populations. To recoup the R&D and production expenses, companies must set high prices.This often leads to negotiations or even rejections by national health systems, which are

reluctant to cover these costs for a small number of patients[23]. For instance, Zynteglo®, a gene therapy for beta thalassemia was withdrawn from the European market due to commercial reasons[183]. Moreover, the complexity of rAAV manufacturing methods and the high doses required per patient also contribute to these elevated costs. The lack of efficient rAAV production methods at large scale, hinders the production of high-quality products in a cost-effective manner.

Considering current manufacturing methods, large manufacturing capacities are needed for approved rAAV products. In addition, the manufacturing demand also depends on the administration method. For example, targeted delivery in the eye as in Luxturna® requires 1.5 x 10^{11} vg/eye for therapeutic expression, whereas Roctavian® for Hemophilia A, administered through intravenous infusion requires 6.0 x 10^{13} vg/kg, so approximately 4.5 x 10^{15} vg per patient need to be injected, skyrocketing the total production scale[184–186] (**Figure 9**). In clinical trials addressing larger patient populations, such as those for Parkinson's disease, targeted delivery strategies have been pursued, while systemic delivery remains feasible only for conditions affecting smaller patient cohorts like Duchene Muscular Dystrophy, where the volume of production would still need to reach millions of liters to treat all EU patients (**Figure 9**).

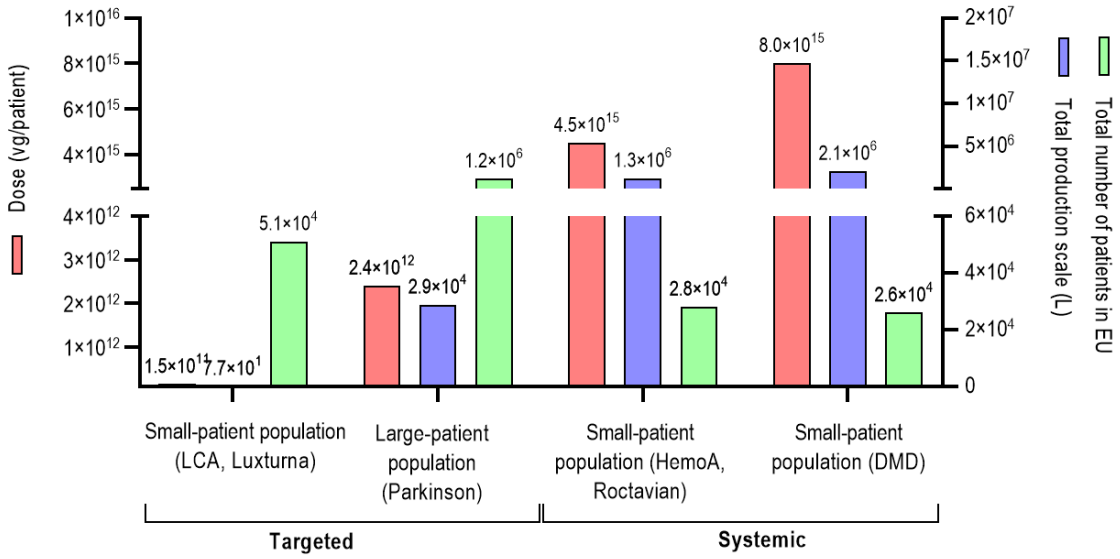


Figure 9. Theoretical manufacturing demand for different rAAV-based therapies based on the dose requirement and patient prevalence in the European Union (EU). Production scale is estimated based on a productivity of 5x10^{14} vg/L and a downstream recovery of 20%. Estimated patient weight of 75 kg in all cases except for DMD, which is estimated as 40 kg. For diseases already addressed by approved rAAV-gene therapy products such as Luxturna and Roctavian, commercial doses are used. Parkinson and DMD rAAV therapy doses are based on NCT00985517 and NCT06114056 from clinicaltrials.gov respectively. DMD, Duchene Muscular Dystrophy. HemoA, Hemophilia. LCA, Leber's congenital Amaurosis. Adapted based on Sha *et al.*[187].

Therefore, the expansion of treatment options beyond rare diseases addressing larger patient populations will pose significant challenges to existing manufacturing systems. The development of cost-efficient approaches that are capable of producing rAAV products with high purity and potency, at higher productivities is key. The advancement of more robust rAAV production methods, which can maintain high yields without compromising vector quality, will not only widen the range of treatment options but also substantially enhance patient access, making gene therapies more broadly available[188].

In this work, the standard rAAV production methods currently employed in the industry are reviewed, with the aim of identifying and understanding the existing production bottlenecks and propose potential modifications to improve them.

4. Strategies for rAAV production

The different strategies that have been developed for rAAV production are designed to mimic the wt AAV replication process. In previous sections, the genetic components required for rAAV production have been described. Additionally, the specific helper virus genes to support rAAV production vary with each rAAV production method, depending on the helper virus system employed. The most used helper virus systems are based on adenovirus, herpesvirus, and baculovirus. Two different approaches are identified based on the gene expression method: transient gene expression (TGE) and stable gene expression (SGE). The TGE can be achieved through gene delivery by either transfection or infection, providing the essential components for rAAV production exogenously to the host cells. In contrast, SGE methods involves the stable integration of some or all necessary elements for rAAV production into the host cell's genome, resulting in the establishment of partial or fully stable cell lines, named packaging or producer cell lines, respectively. Of note, following gene delivery, the pathway of rAAV production is similar to that of wt AAV replication, including protein expression, capsid assembly, genome replication and packaging as shown in **Figure 5**[187].

TGE stands out as a rapid and flexible approach for rAAV manufacturing, particularly suitable for early-stage process optimization due to its flexibility to modify the specific rAAV product facilitating therapeutic transgene screening[189]. In addition, it yields enough quantities for commercial rAAV manufacturing, and it is the approach of choice for manufacturing the approved rAAV products to date. There are several TGE strategies regarding the gene delivery method. One approach involves transient transfection of plasmids carrying the elements for rAAV production into host cells. This

can be achieved using a variety of cationic polymers such as polyethylenimine (PEI) that interacts with the plasmid DNA generating positively charged complexes that are then internalized by the cells[190191]. Alternatively, TGE can be achieved by infecting cells with a wild-type or helper viral vector carrying the helper functions and supplying the required elements for rAAV production (*rep/cap* genes and the GOI). In general, viral vector-based delivery shows higher efficiency and consistency than transfection-based delivery[192].

Despite the TGE advantages, DNA is not integrated into the host cell's genome unlike in SGE approaches, thereby it is lost over time due to cell division and environmental factors limiting production period to typically 72 hpt when harvest is performed[190,193]. In addition, there are several transfection-associated disadvantages. First, coordinated delivery of multiple plasmids into target cells must be achieved and the preparation of these plasmids in high quantities and at high purity levels is required, contributing significantly to the cost of goods of the process. Additionally, important limitations are inherent to the transfection process, such as its variability, which can compromise the robustness of rAAV production and efficiency of transfection during process scale-up[194]. The presence of prokaryotic sequences in the pAAV_GOI plasmid may also rise safety concerns due to reverse packaging[149,195]. The use of infection-based production methods is also affected by several limitations, especially regarding the genetic instability of the viral particles carrying the rep construct[196]. In addition, the viruses, either if wild-type or recombinant versions are used, need to be removed from the final rAAV product in the downstream process. Overall, the variability and lack of robustness associated with TGE approach limit its application at large-scale.

The SGE approach remains the predominant technology for large-scale biopharmaceutical production, providing batch-to-batch consistency[194]. This method, based on the integration of the genes of interest into the host cell genome, avoids the need to generate high amounts of DNA compared to transient transfection method or co-infecting the cells with helper viruses. Moreover, the generated stable cell lines are fully characterized, allowing higher reproducibility than TGE systems that streamline large-scale production which results in lower manufacturing costs. However, the process of generating a stable cell line is laborious and time-consuming compared to the TGE method. In the context of rAAV production, several challenges regarding viral particle toxicity further difficult the development of packaging/producer stable cell line platforms, which have been under research since the late 1990s[58,197,198]. Due to this, the manufacturing of AAVs using SGE has not yet been established at the required levels for clinical applications.

In the following sections, the state of the art of the different developed rAAV production strategies based on TGE and SGE approaches are reviewed, discussing the advantages and limitations of each approach in order to identify further improvements for rAAV production process development.

4.1. Transient gene expression

4.1.1 Production of rAAV by transient transfection in HEK293 cells

TGE in HEK293 through transient triple transfection (TT) is the most widely used strategy for rAAV production[184,199]. This method has been disclosed to be the production method for commercial rAAV-based products such as Zolgensma® and Luxturna® [200,201]. Given the advantages of the TT into HEK293 cells, this is the main approach used in this work.

This production method based on adenoviral helper complementation, was developed by Matsushita *et al.*, in 1998, avoiding the need of AdV helper co-infection by using plasmid co-transfection harboring the adenoviral genes[106]. As mentioned, three plasmids are co-transfected: pGOI_ITR plasmid containing the desired expression cassette with the GOI flanked by ITRs, pRepCap, encoding for the *rep* and *cap* genes of rAAV; pHelper plasmid encoding additional adenoviral helper genes (*E4, E2A* and *VA RNAs*) into HEK293 cells that already stably express the transformed DNA of adenovirus *E1A* and *E1B* genes[181] (**Figure 7B**).

Initially, TT of adherent HEK293 cells for rAAV production was the standard, rendering a system difficult to scale up, highly laborious and variable in terms of yields[151,194]. Adherent cells often depend on animal serum for growth, which has variable composition and is not compatible to GMP standards[194]. Scaling up in adherent systems typically involves increasing the surface area for cell attachment, which can be achieved by cell factory system, HYPERflask or roller bottles as in the case of Luxturna®, in which medium exchange after transfection adds further complexity to the production process[202]. Innovations such as iCELLis fixed bed bioreactor have been developed to enhance the efficiency of large-scale production systems in adherent cultures by allowing the adjustment of cultivation conditions[151,203,204]. Consequently, the transitioning of HEK293 cells adapted to grow in suspension serum-free culture media marked a significant advancement for rAAV production, facilitating a more homogeneous and controllable production environment.

An example of suspension HEK293 cells is the HEK293SF-3F6 cell line, adapted into suspension serum-free culture media in 1998[205]. The Expi293™ suspension cell line developed by ThermoFisher are also commonly used[206]. For suspension-based rAAV production processes,

transfection using polyethyleneimine (PEI, 25-kDa), a cost-effective cationic polymer DNA transfection agent is typically used[207]. Initial experiments by Park *et al.* and Durocher *et al.* using PEI-transfection of serum-free suspension HEK293 achieved yields of approximately 1.4 x 10^4 and 3 x 10^4 vg/cell respectively[208,209]. Subsequent optimizations improved rAAV titers by generating an optimized transfection process and selecting an optimal HEK293-serum free cell line, with the highest yield reported by Grieger *et al.* for different rAAV serotypes[210] and Lieshout *et al*. reaching >10^5 vg/cell[211]. Therefore, outperforming adherent systems and achieving similar yields to other TGE approaches as summarized in **Figure 10**[151,208,210,212]. This suspension systems have been successfully implemented in bioreactors such as stirred tank reactor (STR) and rocking motion bioreactors, operated in batch up to 200 L[210,213].

Significant progress in the design of plasmid systems for rAAV production has been made. Xiao *et al.* in 1998 developed a three-plasmid system for rAAV production with novel modifications both in packaging and helper plasmids. They generated a new packaging plasmid (pXX2) containing an inefficient start codon for the expression of large Rep68 and 78, reducing its expression, and introducing an extra copy of the p5 promoter downstream of the *cap* gene[99]. This approach of relocating the p5 truncated promoter of the *rep/cap* ORF is still used to date as an efficient method to enhance rAAV production[214]. Relevantly, Brimble *et al*. in 2022 demonstrated that placing the RBS of p5 downstream the *cap* gene could lead to the packaging of contaminant sequences within rAAV virions[215]. To mitigate the issue, they modified the p5 promoter (p5-HS) by inserting a DNA spacer between the RBS and the trs, reducing the contamination of rAAV with non-therapeutic DNA[215]. Grimm *et al.* in 1998 also engineered the p5 promoter, replacing it by an MMTV-LTR promoter leading to reduced expression of Rep78/68; however, expression of the VP proteins increased compared with VP levels from standard packaging plasmids, increasing packaging efficiency[216]. Additionally, they developed a two-plasmid production system by combining packaging and helper constructs into a single plasmid (pDG) offering advantages in streamlining the transfection process[216]. This approach has been later followed by Tang *et al*. in 2020 to generate a two-plasmid packaging system, confirming its utility for in-vivo applications[217]. Recently, a study comparing the performance of two different configurations of a dual-plasmid system, identified that the pairing of GOI-ITR with *rep-cap* sequences on one plasmid and isolating the helper functions on another (pOXB), offered the best combination of high productivity, efficient packaging and scalability, outperforming the pDG control[211]. Further approaches such as the optimized Rep-Cap split packaging system developed by Emmerling *et al*. further contributed to optimize Rep protein expression improving rAAV titers[204]. The most recent approach created hybrid

Rep proteins by introducing the 3' end of the AAV2 *rep* gene into other serotype rep constructs, improving both capsid protein expression and genome packaging[49]. Overall, modifications finely tuning the large Rep protein expression impacted positively on rAAV yield and packaging efficiency, while increasing capsid protein expression improved rAAV production, indicating that engineering strategies of the packaging plasmid can lead to optimized rAAV production methods.

In addition, improvements to the helper plasmids have also been performed. Xiao *et al*. initiated the removal of some early and most late genes in the adenoviral helper plasmid (not essential for the rAAV production), generating the pXX6 helper plasmid[99]. This modification not only enhanced rAAV production but also minimized the risk of generating recombinant competent AdV (rcAdV) by homologous recombination with *E1* gene in the HEK293 cell genome[99]. Other work eliminating unnecessary adenoviral sequences still present in pAdHelper plasmids has also been performed resulting in enhanced transfection efficiency[204]. Recently, Polyplus® developed an optimized helper plasmid harboring new genes form other AAV helper viruses (pPLUS® AAV-Helper)[218].

Although current rAAV production systems have been optimized to reduce rcAAV levels, the possibility of rcAAV formation remains through homologous or nonhomologous recombination between the AAV ITR present in the vector and the Rep and Cap sequences present in the packaging plasmid[188,219,220]. rcAAV closely resembles authentic AAV vectors, and these entities cannot be separated by purification process steps. Rather, strategies minimizing the risk of rcAAV formation during vector production have been developed, primarily reducing the length of homologous sequences present in both vector genome and packaging plasmid. To this end, the modifications performed by replacing or altering homologous sequences as in Allen *et al.* 1997[180] and 2000[100]; Grimm *et al.*[216] separating *rep/cap*[180,204] or even introducing introns to generate an oversized *rep/cap* construct[221] have to some extent addressed the issue of generation of rcAAV.

4.1.2 Production of rAAV by infection

4.1.2.1. Herpes simplex virus infection in BHK cells

Herpes simplex virus (HSV) can effectively support wt AAV propagation as a helper virus, thus its helper functions are leveraged in rAAV production processes[91]. This approach has been adopted for large-scale rAAV production, producing sufficient quantities for clinical trials[222,223]. The HSV replication proteins providing essential helper functions for rAAV replication include the helicase/primase complex *UL5, UL8, UL52* and the DNA binding protein ICP8 encoded by the *UL29* gene. Additional HSV proteins that enhance AAV replication efficiency has been identified[224] .

Current rHSV-based rAAV production systems employ a replication deficient rHSV strain lacking ICP27 expression[222]. Therefore, it needs to be propagated in ICP27-expressing Vero cells[225]. The rAAV production involves infecting baby hamster kidney cells (BHK) with two rHSV, one containing the AAV *rep/cap* genes and the other carrying the rAAV transgene flanked by ITRs[223] (**Figure 10).**

Infection of suspension-adapted BHK cells with rHSV has attained high productivity, achieving yields ranging from 8 x 10^4 to 2 x 10^5 vg/cell, and enabling large scale production up to 10 L culture[223,226]. However, several issues are associated with the rHSV technology. First, although a replication-deficient rHSV is used, the purification of the rAAV product from rHSV is needed due to its potential immunogenicity. In addition, the rHSV viral inoculum may experience instability, and its propagation is limited by low-efficiency yields posing operational challenges in rAAV manufacturing[227].

4.1.2.2. Baculovirus-based infection in insect cells

The baculovirus-insect cell expression system (BEV) system uses the baculovirus (BV) ability to infect the Sf9 cell line derived from the *Spodoptera frugiperda* species[228]. The infection of Sf9 cells with BEV has been used for heterologous protein expression since 1980s[229]. Currently, it is used for rAAV production of some early approved commercial products such as Hemgenix®[230], manufactured at UniQure and Roctavian®[231] by BioMarin (see **Figure 8**).

Urabe *et al.* in 2002 first reported the use of the BEV for providing helper functions for rAAV production in insect cells[92]. The initial design required triple co-infection of Sf9 cells with rBV, each delivering one of the three essential genes (*rep*, *cap*, and GOI), named Three-Bac, under strong baculovirus promoters[92]. Although this served as a proof of concept, the palindromic configuration of Rep78 and Rep52 coding sequences resulted in genetic instability of rBVs and loss of its expression after several passages, ultimately resulting in a decreased rAAV yield and limited appplication[92]. Further, the Three-Bac system was unable to produce functional AAV vectors of other serotypes[196,232].

Subsequent engineering strategies partially resolved these limitations. The introduction of the Two-Bac system simplified the production process to a dual rBV infection in which the *rep* and *cap* genes are contained within the same rBV, and GOI in a separate rBV (**Figure 10**). Specifically, Chen *et al.* reported increased genetic stability and higher yields by inserting artificial intron sequences in Rep78, Rep52 and Cap coding sequences[233]. Smith *et al.* further improved BEV stability by modifying the initial codon sequence of Rep78 for a weak codon, facilitating ribosome readthrough

and expression of the downstream Rep52 from the same ORF[234]. Furthermore, it has been observed that sub-optimal expression levels of VP1 result in the generation of non-transducing vector particles[196], while strategies restoring VP1 expression resulted in improved vector function[233,235].Recent advancements integrating all the AAV components in a single rBV have further simplified the production process requiring only a single infection[236].The BEV/Sf9 system achieved by TGE has benefited from all the molecular optimizations of rBV vectors, as reviewed by Joshi *et al.*[237] and is now considered an interesting option for rAAV production. Among its advantages, the ability of Sf9 cell line to grow to grow in chemically defined serum-free culture media facilitated lot-to-lot consistency of Sf9 suspension culture, enabling the production of high rAAV amounts though scalable processes. The Three-Bac platform has been used at 200 L scale, and up to 2 x 10^4 vg/cell have been achieved using the Two-Bac platform[196,238]. However, although BVs are non-pathogenic to humans[239], its removal from the final product is still required[240]. Despite these improvements, the BEV/Sf9 system exhibits several quality limitations compared to the transient TT system in HEK293 cells. These challenges are discussed in subsequent sections.

4.1.3 Comparison of the two most used rAAV production platforms: BEV/Sf9 and TT in HEK293 cells

For the production of rAAV-based gene therapy treatments approved for clinical use, both the TGE approaches based on the BEV/Sf9 system and TT into HEK293 cells are used. Consequently, a comparative analysis of these two systems, in terms of quantity and quality of the produced rAAV is essential to provide critical insights into the bottlenecks and potential advantages of each system.

The initial development of BEV/Sf9 rAAV production system was driven by the difficulties in scaling-up and limited yields observed with early systems based on adherent HEK293 cells. To date, improved suspension serum-free rAAV production systems in HEK293 cells have been reported reaching 10^5 vg/cell[210], which are comparable to the yields achieved by the Two-Bac BEV/Sf9 system at 7.8 x 10^4 vg/cell[234,237,241]. Narrowing the differences in rAAV yields, the quality differences between the two rAAV production systems remains a key attribute into determining the preference for one system over the other. The quality concerns associated include the following.

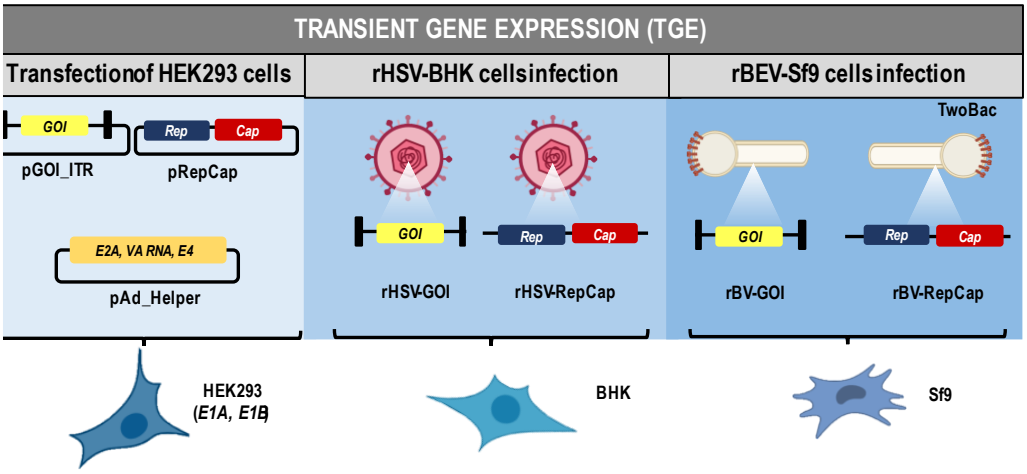
First, rAAV vectors produced using the BEV/Sf9 system have shown lower infectivity levels than those produced in human cell lines[242], potentially requiring higher doses for patient treatment. This increase in dose not only could outweigh any improvement in productivity achieved with the BEV/Sf9 system, but also may have implications in clinical use, such as vector toxicity at high doses[240]. As described, this poor transduction efficiency has been attributed to issues with VP1

expression and incorrect VP stoichiometry, as when optimizations to VP ratios are made, greater transduction is achieved[233,235].

Second, differential post-translational modifications (PTM) in rAAV capsids derived from insect cells compared to those derived from HEK293 cells have been identified[240]. Differential PTM may influence host cell proteins (HCPs) as impurities of the final product, where N-glycan PTMs found in insect cell impurities could trigger allergic reactions in humans[243]. The impact of these differential PTMs on receptor binding affinity and potential effects on potency remains unclear to date.

Third, the BEV/Sf9 system has been reported to produce a high degree of truncated and unresolved rAAV genome species compared to TT-HEK293 based systems[244]. These genome variations may affect the rAAV performance. In addition, rAAV genomes from Sf9 cells have been found to be more methylated in key repressive regions than those from HEK293 cells[240].

Given the outlined quality and quantity differences between these two production approaches, this work is focused on the optimization of rAAV production using the human cell platform[210,212,245].



Yield	> 10^5 vg/cell[(a)]	$5x10^4$ vg/cell to $1.5x10^5$ vg/cell[(b)]	$7.8×10^4$ vg/cell[(c)]
Safety concerns	Generation of rcAAV by plasmid recombination which may be reduced by plasmid design	Helper virus contamination Cell line of non-human origin	Cells of non-mammalian origin Immunogenic reactions from impurities derived from insect cells
Advantages	Simplicity and flexibility of rAAV production. Helper virus free	Suspension scale up culture is suited	Suspension scale up culture is suited
Challenges	Variability and low scalability of triple transfection	Production of two rHSV helper viruses in VP27 Vero cell line rHSV stability	rBV stability during expansion Ratio of VP may be different from wt AAV capsids rAAV potency less than from HEK293-derived cells
References	Grieger *et al.* (2016)[210] Chahal *et al.* (2014)[207] Park *et al.* (2005)[209] Durocher *et al.* (2007)[208] Lieshout *et al.* (2023)[211]	Thomas *et al.* (2009)[2] Knop *et al.* (2011)[226]	Urabe *et al.* (2006)[196] Chen *et al.* (2003)[233] Smith *et al.* (2009)[234]

Figure 10. rAAV manufacturing methods by the transient gene expression (TGE) approach.[(a)] Highest specific production reported by Grieger *et al.* (2016)[210] and Lieshout *et al.* (2023)[211].[(b)]Highest reported by Knop *et al.* (2011)[226]. [(c)]Highest reported by Smith *et al.* (2009)[234].

4.2. Stable gene expression

In rAAV production, the generation of stable cell lines for SGE can reduce the complexity and variability associated with TGE systems, thereby enhancing control over the manufacturing process[246]. To this end, various rAAV packaging and producer cell lines have been developed using different strategies[188,247]. Packaging cell lines represent a first step towards establishing producer cell lines, in which some of the necessary elements for rAAV production have been integrated, but still require transfection or helper virus infection to supply all the required elements for rAAV generation. For this reason, efforts have been made on the development of a completely stable producer cell lines integrating all elements for rAAV production: GOI, *rep* and c*ap* genes, and also the necessary viral helper genes[248]. This strategy eliminates the need of additional transfection or helper virus infection[249]. The advancement of helper virus-free, stable producer cell lines has received increasing interest as a promising approach. Choosing the right host cell is crucial for the success of both packaging and stable producer cell line development, as discussed in the following sections[250–254].

4.2.1 rAAV packaging cell lines

4.2.1.1. E1- transformed cells: HEK293

To date, several efforts to establish rAAV packaging cell lines based on E1-transformed cells, such as HEK293 cells have been made, presenting critical challenges[197,255,256]. When HEK293 is used, the main limitation is the cytotoxicity driven by *E1A*-mediated transactivation of AAV *rep*[58,197] and adenoviral helper[257] genes. Rep products are cytostatic and cytotoxic to the host cells[58,197]. In addition, some adenovirus helper genes such as E4 products can induce apoptosis[198]. To address this issues, strategies to tightly regulate Rep and/or AdV helper expression have been widely investigated[197,246,255,256,258].

Initial strategies focused on regulating or substituting the p5 promoter driving the transcription of large Rep protein with inducible promoters such as the glucocorticoid-responsive mouse metallothionein transcription promoter (MMTV) or doxycycline-inducible promoter[197,258,259]. However, controlling only p5 products could not eliminate the growth disadvantage on cells derived from leaky expression of p19 products, revealing that expression of small Rep protein was still negatively affecting cell proliferation[197,255,258,259]. However, while replacing p5 promoter for an

inducible one is feasible[197], this is a non-straightforward task for p19 products as it is contained in the coding sequence of the large Rep proteins[48].

To address this, an innovative approach named “dual splicing switch” was developed by Qiao *et al.* to regulate all Rep proteins[256]. This strategy involves inserting an intron flanked by transcription termination sequences into the shared coding region of the Rep proteins, prematurely ending their transcription while maintaining p5 and p19 endogenous promoters. Although plasmid transfection is not required, the infection of an E1-deleted AdV carrying Cre-recombinase is required to mediate the excision of the termination sequence and reactivate the *rep* genes becoming a by-product to be removed[256,260] (**Figure 11**). Moreover, the E1 region into the HEK293 cell line raises the possibility of generating rcAdV. Further, the reliance on adherent cultures of this packaging cell line adds complexity and scale-up limitations to this system. While transition to suspension cultures should be feasible, this has not been yet performed.

4.2.1.2. Non-E1 transformed cells: A549 and HeLa

To circumvent the challenges associated with the continuous expression of the AdV *E1A* gene in HEK293 cells, several efforts have been directed towards the generation of *rep/cap* stable packaging cell lines using non-E1 transformed cells such as A549 and HeLa cells (summarized in **Figure 11**). The mechanism driving rAAV production in these packaging cell lines involves the amplification of *rep/cap* sequences enabling the production of sufficient Rep and Cap proteins for rAAV genome replication[261,262]. This episomal amplification of integrated *rep/cap* DNA sequence is dependent on the cis-acting replication element (CARE) mapped at the p5 promoter region of the *rep* gene (nt 190 to 540 of AAV2) and requires the activity of the AdV DBP, cellular polymerases and Rep proteins[262,263]. Chadeuf *et al.* demonstrated that CARE amplification, essential for rAAV production, is efficiently triggered by wt AdV infection and not by plasmid transfection containing the complete adenoviral genome[252]. Moreover, the CARE-dependent amplification process has been further elucidated to require the AdV MLP activity, specifically L4 22/33K protein expression[264].

The initial generation of packaging cell lines was developed by Clark *et al.* in 1995, using the HeLa cell line which lacks the *E1* gene through the stable integration of *rep/cap* genes and triggering rAAV production by wt AdV infection[253]. Subsequent developments have shown the scalability of this approach using the HeLaS3 clone capable of growing in serum-free suspension cultures reaching up to 250 L[250–253,261,265]. An alternative method employed the HeLa cell lines for rAAV production using an Ad-AAV hybrid method, consisting on using two AdV helper viruses: first a

helper wt AdV with functional *E1* genes to induce *rep/cap* expression, followed by a replication defective Ad-AAV virus, in which the rAAV GOI replaces the E1 region sequence[251,266].

Notably, Qiao *et al*. explored a novel strategy towards the development of a rAAV producing cell lines using HeLa cells[257]. This method, differs from other approaches as it involves the stable integration of the *E1A* gene driven by a tetracycline-responsible promoter via the Tet-Off system, alongside the *E1B* gene under its endogenous promoter and the GOI[257]. Although E1A inducibility potential over time decreased, this method evidenced the challenges associated with a complete stable producer cell line, as the stable integration of the AdV helper genes proved unfeasible, resulting in a cell line that still required replication-defective AdV co-infection[257].

The A549 cell line, which lacks AdV gene expression, was also investigated for rAAV production[248,254]. The feasibility of A549 for the rAAV production by generating packaging cell lines demonstrated that the HPV sequences contained in HeLa cells were not necessary to drive rAAV production nor enhance it, as even higher *rep/cap* amplification was observed in A549-packaging cell lines over wt AdV infection. In the most recent approach, A549 *rep/cap* packaging cell line was generated and production of rAAV was accomplished by infecting the packaging cell line with a temperature-sensitive AdV[248]. This enables transition of HeLa to cells such as A549 devoid of *E6* and *E7* HPV genes, presenting a safer alternative reducing the risk of oncogenic DNA encapsulation in the rAAV capsid[267].

Despite this advances, a primary limitation when using non-E1 transformed cells for SGE systems is the need of wt AdV of tsAdV infection to trigger rAAV production, which presence in the final product is highly undesirable for vector safety, and removal is necessary to avoid potential immune response and inflammatory response effects[268].

4.2.1.3. BEV/Sf9-based packaging cell lines

The development of packaging cell lines for rAAV production has also been performed using the BEV/Sf9 technology[237]. This approach, named OneBac, is based on the stable integration of *rep* and *cap* genes into the Sf9 cell genome, reducing the rBV infection requirement to a single vector with the GOI (**Figure 11**). In the OneBac system, the expression of *rep/cap* genes remains transcriptionally inactive until triggered by rBV infection, which initiates the amplification of the integrated genes, enhancing their expression[241]. This is facilitated by the immediate-early trans-regulator 1 (IE-1) protein, provided by the rBV infection, which activates the IE-1 binding target sequence hr2-0.9 within the integrated sequences boosting its expression[241]. The OneBac system

offers scalability benefits, given the ability to easily expand Sf9 cells in suspension, and has been developed for a variety of AAV serotypes. In addition, genetic modifications addressing suboptimal VP1 protein expression levels enhancing rAAV yield has also been performed[269–273]. Nonetheless, the Two-Bac system has been more extensively used for large-scale rAAV production, offering a simpler and more flexible system compared to the OneBac system[237].

4.2.2 rAAV producer cell lines

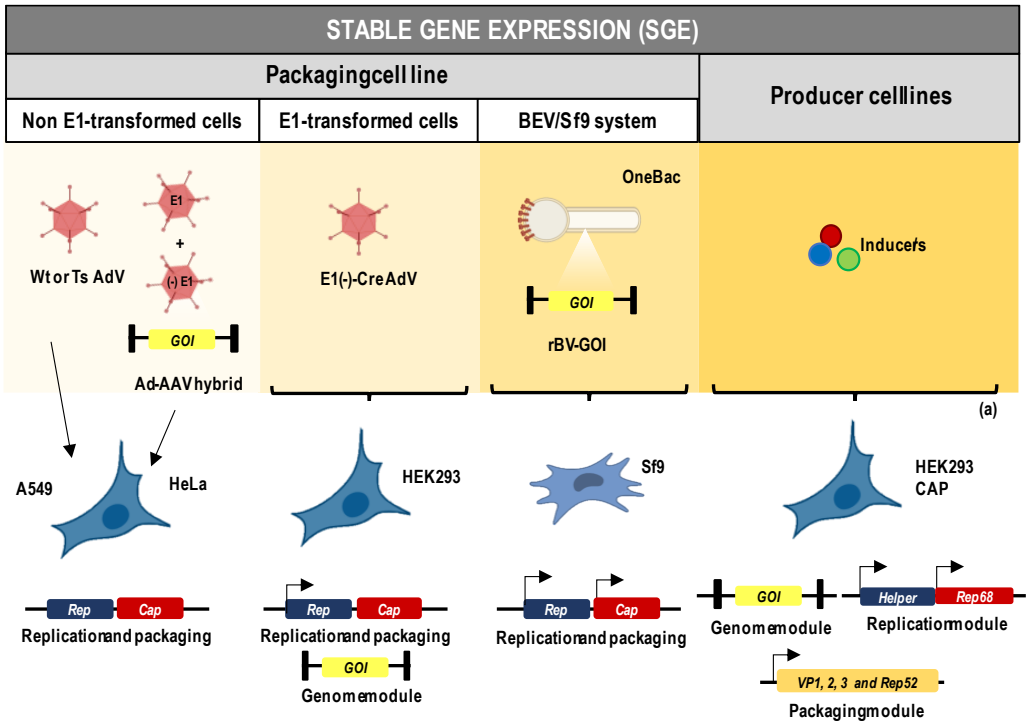
Recently, several successful attempts have achieved the generation of a completely stable rAAV producer cell line, helper virus-free, using inducible systems in E1-transformed cell lines[274,275].

In one approach, rAAV producer cell line using CEVEC's Amniocyte Production (CAP) cells named ELEVECTA[274] was accomplished through stable integration of the *rep* gene, AdV *E2A* and *E4orf6* helper genes and *cap* gene; all controlled by the Tet-inducible system, alongside the integration of the GOI[274]. This approach allows scalable rAAV production by induction, eliminating the need for helper virus infection or additional transfections. Nonetheless, a potential limitation arises from the coordinated expression of all genetic components under the same inducible system, which may lack the capacity of fine-tuning individual viral expression levels.

Another approach of a transfection-free, helper virus-free rAAV producer cell line in HEK293 cells has been developed using a synthetic biology approach by Lee *et al.*[275]. This development enables independent control over replication and packaging modules using multiple inducible systems. Destabilizing domains (DD) gene-inducible system and Tet inducible system are used to regulate Rep68 expression both at transcriptional and post-translational levels. The GeneSwitch promoter regulates *E2A* and *E4orf6* gene expression, while VP1, 2, 3 and Rep52 are expressed under the regulation of the CumateSwitch control[275]. Although this approach successfully allowed control over rAAV titer and full particle content, the productivity of the generated stable cell line was still two orders of magnitude lower compared to the triple transfection-based approach[275] (**Figure 11**).

Although various attempts have been made to generate rAAV producer cells both in E1-containing and E1-deficient cell lines, to the best of our knowledge, no successful approach has achieved high-yield rAAV production comparable to the standard TT into HEK293 cells, and that is free from helper virus infection. Moreover, the packaging and producer cell lines generated to date involve the stable introduction of a specific capsid serotype and GOI, limiting its flexibility and requiring the generation of a novel cell line for every rAAV product. This process is time-consuming

and evidences the need for the development of more flexible and efficient cell lines, capable of producing high yield and quality rAAVs.



Yield	$>10^4$ (b)	$1.3x10^5$ vg/cell	up to $2.16x10^5$ vg/cell (c)	ND. for ELEVECTA® $>10^3$ vg/cell Pf3 by Lee *et al.* 2022
Safety concerns	wt AdV or Ts AdV infection required	rAdV-Cre vector infection	Cells of non-mammalian origin Immunogenic reactions from impurities derived from insect cells	Antibiotic resistance procaryote sequences stably integrated
Advanta-ges	No transfection is required	All Rep proteins controlled by dual-splicing system	Suspension scale up culture is suited Reduction from two to one rBV infection.	Control over product quantity and quality in Lee *et al.* approach Eliminate the need of infection or transfection
Challenges	Co-production of wt AdV	Generation of rcAdV Adherent culture requiring serum	rBV stability during expansion Ratio of VP may be diferent from wt AAV capsids rAAV potency less than from HEK293-derived cells	Low titer Fixed serotype and GOI
References	Farson *et al.* (2004)[248] Gao *et al.* (2002)[254] Liu *et al.* (2000)[261] Clark *et al.* (1995)[253] Gao *et al.* (1998)[251]	Qiao, Wang, *et al.* (2002)[256]	Joshi *et al.* (2019)[272] Wu *et al.* (2019)[273] Aslandini *et al.* (2009)[241] Mietzsch *et al.* (2014)[271]	Coronel *et al.* (2021)[274] Lee *et al.* (2022)[275]

Figure 11. rAAV production approaches to stable gene expression (SGE). (a) Schematic representation from Lee *et al.*[275]. (b) Maximal specific productivity reported by Farson *et al.* (2004)[248] (c) Maximal specific productivity reported by Wu *et al.* (2019)[273].

5. Operation systems for rAAV production through triple transfection into suspension HEK293 cells

As mentioned, currently, the TGE method via TT in suspension HEK293 cells is the workhorse for rAAV production, enabling rapid and flexible generation of rAAV at relatively high quantity and quality. However, to meet increased production demands, the scale up of this process requires to address several challenges to maintain a high productivity. Various operation systems can be used to maximize rAAV productivity in suspension HEK293 cell lines.

5.1. Batch cultures

In the initial stages of TGE development, small scale production systems facilitate high-throughput analysis for optimization purposes. Suspension HEK293 cells are typically cultured in shaken polycarbonate flasks, which are placed on orbital shakers under controlled conditions. Smith *et al*. in 2003 first reported rAAV production by TT into HEK293 cells in suspension conditions using Erlenmeyer flasks[276]. For scaling up production, large shake flasks offer a straightforward option, with reported working volumes ranging from 200 mL to 1 L[190]. When the demand for higher production volumes increases, those can be obtained using bioreactors such as stirred tank reactor (STR) and rocking motion bioreactors, operated in batch mode[208–210,277]. For rAAV production for preclinical development, scale-up has been successfully implemented for volumes up to 200 L[213]. In addition, scalability between 50 and 500 L production using STR single use bioreactors has been demonstrated by PALL®[278].

In batch culture, cells are inoculated at low cell density and transfected once the viable cell density reaches the optimal level for transfection, typically around 1 – 2 x 10^6 cells/mL[207,210]. Due to the transient nature of the gene delivery, processes are often limited to 72 hpt, after which the cells are subjected to lysis for full rAAV recovery[210,279].

Noteworthy, several limitations inherent to the TT in suspension HEK293 cells process present challenges for large-scale production. First, the TT protocol is subjected to a high variability, and alterations in the transfection process can affect the transfection efficiency, causing lot-to-lot variability and introducing heterogeneity in the process during scale-up[280]. In addition, a decrease in cell-specific productivity is observed when cells are infected or transfected at high cell density cultures (HCD), known as the cell density effect (CDE)[190,281]. This phenomenon, which is cell line and culture media-dependent, has been shown to significantly reduce rAAV production by TT into

HEK293 cells in batch[282–286]. Therefore, the scale-up of batch operation mode is limited by cell density at the time of transfection, limiting volumetric productivity[287].

5.2. Fed-batch cultures

Fed-batch culture strategies offer a significant improvement over batch cultures. In fed-batch, fresh media is added at discrete time points and exhibit higher growth rate and prolonged cell growth phase compared to batch. This is largely attributed to the dilution of inhibitory compounds that accumulate during cell culture negatively affecting cell health and productivity. As a result, cells can reach higher densities and maintain their productivity over an extended period, leading to a higher rAAV yield[288]. While this process has been used for rAAV production using the BEV/Sf9 production method by TGE[272,288], its use in TT into HEK293 cell cultures has been more limited. Fujifilm Irvine Scientific reported the development of feed prototypes, which added to the cell culture 24 hpt improved rAAV production and virus packaging efficiency in 30 mL shake flask[289]. However, no higher-scales of fed-batch for rAAV production by TT into HEK23 cells has been reported to date.

5.3. Continuous culture strategies: perfusion operation mode

Although batch and fed-batch culture modes are preferred in industry for recombinant proteins and monoclonal antibodies, continuous culture strategies have demonstrated to streamline bioprocessing costs for viral vector production and viral particle production, such as lentivirus, retrovirus, adenovirus and AAVs[210,286,290–292]. Continuous intensification strategies, such as perfusion culture, are based on retaining cells at the bioreactor, using cell retention devices, continuously providing fresh media while removing spent media to maintain a constant reactor volume (**Figure 12**). For rAAV production, the most commonly used cell retention devices are the Alternating Tangential Flow (ATF), Tangential Flow Filtration (TFF), and Tangential Flow Depth Filtration (TFDF)[286,293]. Perfusion culture mode has been successfully implemented to increase viral vector productivities and cell densities, improving rAAV manufacturing through either TGE or SGE using a double approach.

First, media replacement strategies have been described as a mean to modulate the cell physiological condition associated with the CDE, proving valuable for achieving effective transient transfection at HCD[281,286]. In this regard, Mendes *et al.* observed a 5-fold reduction in rAAV specific production when cells were transfected at a HCD of 5 x 10^6 cells/mL compared to 2 x 10^6 cells/mL using the batch operation mode[286]. However, when a medium exchange was performed before

transfection, a threefold increase in rAAV titers was achieved, demonstrating a clear correlation between medium exchange and enhanced rAAV production[286].

Second, the use of perfusion mode of operation can also be used to exploit the secreted rAAVs, by harvesting the cell culture media[151,152,210,286]. Harvesting only the supernatant fraction avoids introducing additional contaminants compared to rAAV harvest by lysis in batch production, facilitating subsequent purification steps with much fewer cell-derived materials[52,153,157].

The use of perfusion operation mode to exploit the rAAV egress property was first performed by Grieger *et al.* in 2016[210]. At small scale, medium exchange was conducted by cell culture centrifugation every 24 hours, resulting in a 4.8 − 6.5-fold increase in rAAV8 and rAAV9 yield respectively, compared to batch control[210]. This was scaled up in a WAVE bioreactor using a perfusion set up, demonstrating that continuous media exchange allowed the continuous harvest of the rAAV from the cell culture media with minimal loss of the remaining rAAV at the cell pellet fraction[210]. Recently, the use of ATF and TFDF devices for rAAV recovery at the extracellular fraction have been compared during perfusion culture, with TFDF enabling a 100% recovery[286].

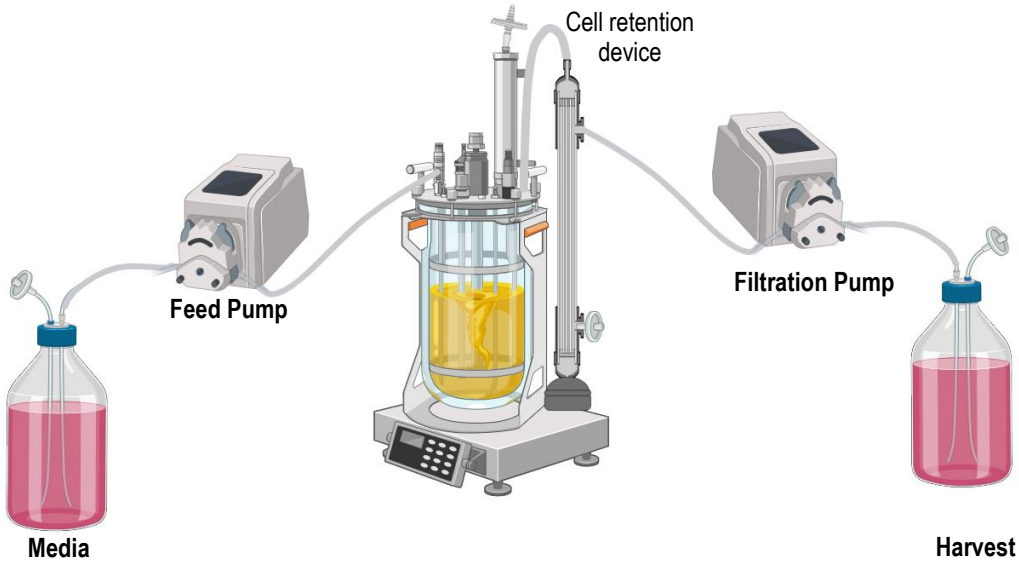


Figure 12. Schematic representation of large-scale rAAV manufacturing using perfusion system. Continuous production mode by perfusion, where continuous media exchange enables continuous harvest from the cell supernatant fraction. The implementation of the EGE method by re-transfecting at discrete time points is a promising strategy to enhance rAAV yields.

5.4. The extended gene expression protocol

The Extended Gene Expression (EGE) protocol, developed by Cervera *et al.,* has emerged as a promising tool for enhancing viral particle production[214,294]. This protocol, developed for the production of VLPs (virus like particles) involves repeated medium exchanges combined with multiple rounds of transfection, leading to a 4 − 12-fold increase in Gag-GFP VLP production in HEK293 cells[294]. This approach combines the flexibility of TGE with an extended production timeframe, introducing more functional DNA, from typically 72 hpt harvest in TGE, up to 240 hours[214,294]. The implementation of the EGE protocol for the production of rAAV has resulted in a remarkable threefold increase in total rAAV harvested from the supernatant, compared to TT production in batch and cell lysate recovery[214]. Viral particle production using the EGE method adapted to bioreactor scale using perfusion mode of operation has also been successfully demonstrated both for VLPs and rAAVs[295]. The progress in developing and implementing the EGE methodology at large-scale using perfusion culture mode allows for the intensification of rAAV production strategies positions this method as a promising manufacturing strategy to overcome the challenges faced by rAAV-manufacturing processes.

However, the EGE method requires a high amount of GMP-grade plasmids for repeated transfection rounds, which substantially increases the costs of the process. Overall, the described rAAV production systems still suffer from high complexity and variability, specially in the TGE approaches, and low titer and flexibility in the SGE developed to date. This underscores the need for the development of new production processes and improving the existing ones in terms of both yield and quality, while ensuring a cost-efficiency. These properties must be robustly preserved during scale-up. In this thesis, steps have been taken in this direction, paving the way for the broader application of gene therapies in clinical settings.

OBJECTIVES

The general objective of this study is to identify and improve current rAAV production methods and to establish the groundwork for the development of new rAAV production methodologies using stable producer cell lines. Specifically, it is proposed to:

1. Analyze molecular changes in HEK293 cells resulting from the transient transfection method of rAAV production in comparison to those from the extended gene expression method.

2. Identify potential bottlenecks affecting rAAV production by transient transfection and explore modulators to improve it.

3. Enhance the rAAV production at the extracellular fraction by metabolic engineering of HEK293 cells.

4. Mitigate the cell density effect in HEK293 cells to improve the scalability of the transient transfection method.

5. Characterize the *E1* gene expression in HEK293 cells during rAAV production and develop methods to maintain its expression.

6. Assess the feasibility of generating stable rAAV producer cell lines in HEK293 cells using inducible *rep* gene expression.

7. Evaluate the potential of the suspension, serum-free A549 cell line for generating stable rAAV producer cell lines through inducible *E1A* gene expression.

RESULTS

PART I. Enhancement of rAAV production through transient transfection

CHAPTER ONE

Quantitative proteomic analysis of HEK293 provides insights of host cell response associated to the extended gene expression method of rAAV production

INTRODUCTION

Recombinant adeno-associated viruses (rAAV) have emerged as one of the most promising vectors for gene therapy applications. This is attributed to their ability to elicit robust and long-lasting transgene expression, lack of pathogenicity of its wild-type (wt) infection, low immunogenicity, and the ability to transduce both dividing and non-dividing cells[40,188]. Recently, several rAAV-based gene therapies have been approved by the European Medicines Agency (EMA) for clinical use[3]. However, a major challenge limiting its broader application is the development of efficient manufacturing systems capable of producing the high yields required for human rAAV gene therapy at an affordable price.

The most commonly employed method for rAAV production is transient gene expression (TGE) via triple transfection (TT) into HEK293 cells[277,293]. This approach has been disclosed to be used as the production method for Luxturna® and Zolgensma®[200,201]. In TT, typically three plasmids are co-transfected: one encoding the therapeutic gene of interest (GOI) flanked by AAV2 ITRs (pGOI), one encoding the AAV *rep* and *cap* genes (pRepCap), and one with adenoviral (AdV) helper genes *E4, E2A* and *VA RNAs* (pHelper). In addition, HEK293 cells harbor the AdV5 helper genes *E1A* and *E1B*[181]. Proteins encoded by the *E1A* gene are required to initiate rAAV replication and modulate the cellular environment. Originally, adherent HEK293 cells were the standard for rAAV production, rendering a system difficult to scale up and highly variable in terms of yields[151,194]. Consequently, the development of HEK293 cells able to grow in suspension serum-free culture enabled the transition into large-scale rAAV production, suitable for cGMP standards and achieving comparable yields to those obtained in adherent cell cultures[207,208,210,212,296].

The TGE technology for rAAV production stands out as a versatile and flexible approach, as the pGOI and the pRepCap plasmids can be easily modified depending on the specific rAAV product. Further, by splitting elements across different plasmids, the optimal ratio can be finely optimized to maximize rAAV production[204]. Despite the benefits of TGE, several challenges limit its application for rAAV production at large scale. First, coordinated delivery of multiple plasmids into target cells must be achieved and the preparation of these plasmids in high quantities and at high purity levels is required, contributing significantly to the cost of goods of the process. Moreover, plasmid DNA remains episomal into transfected cells, limiting production timeframe. Additionally, important limitations are inherent to the transfection process, such as its variability, which can compromise the robustness of rAAV production and efficiency of transfection. Transfection efficiency is

significantly reduced in high cell density cultures (HCD), known as the cell density effect (CDE)[190,281]. This phenomena has been demonstrated to affect viral vector production[282–286]. This is important when implementing TGE to large-scale production, where higher cell densities can be achieved[287]. Thus, unravelling factors governing transient transfection efficiency, especially at HCD may favor the exploitation of TGE at large-scale, contributing to achieve higher rAAV titers. While TGE allows for readily rAAV production, these outlined challenges highlight the demand for the development of more efficient production processes.

In contrast, the extended gene expression (EGE) protocol was developed by our group as a promising tool to enhance viral particle production[294]. The EGE production strategy consists in performing consecutive medium exchanges (ME) every 48 hours, to remove metabolic waste and renew fresh culture media with the combination of repeated rounds of transfection (RT) to achieve a sustained level of gene expression to expand the production phase[214]. The EGE methodology has been successfully used for rAAV9 production enhancement, increasing by 3-fold the total rAAV harvested from the supernatant, compared to TT in batch and cell lysate recovery[214]. Thereby, enabling the continuous harvest of the secreted rAAV fraction from the supernatant of the cell culture only, eliminating the need of lysing the cells for rAAV harvest[214,295,297]. This facilitates downstream processing steps also contributing to sensibly reducing host cell protein and DNA contaminations[214]. Viral particle production using the EGE method adapted to bioreactor scale using perfusion mode of operation has also been demonstrated[295]. The use of cell retention devices coupled with constant addition of cell media ensured sustained cell culture homeostasis throughout the culture period. The implementation of perfusion mode of operation has also been related to the modulation of CDE, enabling re-transfection by modulation of the cell metabolic and physiological state[285,298]. The progress in the development of the EGE methodology position it as a promising strategy to overcome the challenges facing rAAV-manufacturing processes. This method extends the short production times associated with the TGE approach, while it maintains the product flexibility of this system. Moreover, being compatible with large-scale continuous manufacturing processes enables a competitive alternative to SGE platforms for producing rAAV vectors on a larger scale.

However, little is known about the molecular mechanisms by which the EGE protocol enhances rAAV production in HEK293 cells. The rAAV production is an intricate dynamic between viral protein and host cell proteins, in which the viral protein hijacks the host cell for its replication. Various approaches based on transcriptomic and proteomic analysis have been made towards characterizing biological pathways enriched during TT method of rAAV production in HEK293

cells[277,299–301]. Although global biological pathways enriched during rAAV production were detected, limited enhancement of rAAV production method was provided.

In this study, the dynamics of the host cell proteome during rAAV production in HEK293 cells are investigated. A multiplexed quantitative proteomic analysis was conducted to various rAAV production protocols compared to cell growth at different time points. Two methods of rAAV production are compared: the standard transient TT and the EGE protocol. In the study, the analysis of differential protein and enriched molecular processes provides understanding into each cellular state. Gaining this insight helps detecting production and transfection bottlenecks affecting the TT method of rAAV production and host cell physiology changes over the EGE implementation. These findings may provide new keys for further rational metabolic engineering of host mammalian cell lines to support high-yield rAAV production.

MATERIALS AND METHODS

Mammalian cell line and culture conditions

The cell line used in this work is a serum-free suspension-adapted HEK293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from the National Research Council of Canada (NRC, Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL flasks with vent cap (Corning, NY, USA) at 37ºC, 5% of CO2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland).

Cell culture media was FreeStyle™ F17 Expression media (Gibco, ThermoFisher Scientific, San Jose, CA, USA) supplemented with 8 mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA), 0.1% Pluronic™ F-68 Non-ionic surfactant (Gibco, ThermoFisher Scientific, San Jose, CA, USA) and IGF-1 (91590C, Sigma-Aldrich, San Luis, MO, USA) at a final concentration of 50 μg/L. Cell concentration and viability were determined using the NucleoCounter® NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

Plasmids for rAAV9 production

Plasmids used for triple transfection were pHelper, pGOI and pRep2Cap9 kindly provided by Dr. Fatima Bosch from Centre for Animal Biotechnology and Gene Therapy (CBATEG), at Universitat Autònoma de Barcelona (UAB, Bellaterra, Spain). Plasmid pRep2Cap9 encodes the Rep proteins from AAV2, and the Cap proteins from AAV9. Plasmid pGOI (pcohSgsh) encodes for a therapeutic gene expression[214]. The pHelper plasmid codifies for the adenovirus functions necessary for rAAV replication and packaging, namely: *E2A*, *E4* and VA RNA I and VA RNA II. Plasmids were produced at Plasmid Factory and stored at -20ºC (Bielefeld, Germany).

Transient transfection for rAAV9 production

Transfections were carried out at a cell density of 2 x 10^6 cells/mL using plasmid DNA ratios and concentrations previously described for rAAV9 production in Erlenmeyer flasks using 10 mL of total culture volume[214]. Polyplexes were formed by adding PEI to plasmid DNA diluted in fresh culture media (5% of the total culture volume to be transfected). Transfection reagent PEIpro® (Polyplus, Illkirch-Graffenstaden, France) was used. Briefly, the corresponding DNA mixture was diluted with supplemented F17™ media and vortexed for 10 seconds. Then PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, incubated for 15 minutes at room temperature and then added to the cell culture.

For rAAV production by the TT method, cell culture was harvested at 72 hpt or 96 hpt. For EGE method, media exchange (ME) was performed at 48 hpt by centrifuging the cell culture at 300 x g for 5 minutes. The cell supernatant was harvested, and the cell pellet was gently resuspended with fresh media. After that, re-transfection (RT) was performed according to previous transient transfection methodology. All conditions were performed in triplicate.

For rAAV harvest for proteomic analysis, cell culture samples were taken at 48, 72, and 96 hpt by centrifuging the culture at 1000 x g for 10 minutes. The supernatant was collected, and cellular pellets were stored. For cell lysate collection, cells were treated with 1x cell lysis solution (500 mM Tris, 1% Triton X-100 (v/v), 20 mM $MgCl_2$ at pH 7.5) and incubated at 37ºC, 130 rpm, for 1 hour. Lysate mixture was centrifuged at 4000 x g for 20 minutes and supernatant was collected as cell lysate fraction. All fractions were stored at -80ºC until analysis was performed.

Quantitative polymerase chain reaction (qPCR) for rAAV quantification

For the determination of physical vector genome titer of rAAV (vg), 5 µL of the cell lysate fraction or cell supernatant fraction from each sample was added to DNase I (0.066 mg/mL, 10104159001, Roche, Basel, Switzerland) prepared in DNase I buffer (0.16 mM $CaCl_2$, 6.6 mM $MgCl_2$, 17.3 mM Tris-HCl, 0.066 % Pluronic F68, pH 8.0). This mixture was incubated at 37 ºC for 16 hours. Then, 4 µL of 30 mM EDTA (AM9260G, Invitrogen, Carlsbad, CA, USA in 10 mM Tris-HCl, pH 8.0) was added to each tube and incubated at 70ºC for 30 minutes to inactivate DNase I. After, 5 µL of Proteinase K (2 mg/mL, EO0491, ThermoFisher Scientific, San Jose, CA, USA) was added, incubated at 55ºC for 2 hours and inactivated at 95ºC for 15 minutes. The quantification of viral genomes was conducted by quantitative PCR (qPCR) as previously described[214]. This reaction was set up using iTaq Universal SYBR Green Supermix (1725124, BioRad, Hercules, CA, USA) and a set of primers targeting the rabbit beta-globin polyA sequence present in the expression cassette. Thermocycling profile was based on manufacturer's recommendations in a CFX384 Touch Real-Time PCR Detection System.

To quantify the accumulated rAAV titer in the supernatant (SN) (extracellular media), **Equation (1)** was used. To quantify the rAAV accumulated titer in the whole cell culture, the cell lysate (Lys) fraction, **Equation (2)** was used. The total accumulated rAAV in the supernatant (Acc. rAAV in SN) was determined by adding the titer of supernatant fraction at harvest (96 hpt) to the titer of the harvested supernatant fraction 48 hpt. Similarly, the total accumulated rAAV in the cell lysate fraction (Acc. rAAV in Lys) was determined by adding the titer of cell lysate fraction at harvest (96 hpt) to the titer of the harvested supernatant fraction 48 hpt.

$$\text{Acc. rAAV in SN (vg)} = \text{rAAV titer in SN 96 hpt} \times V + \text{rAAV titer in SN 48 hpt} \times V \quad \textbf{(1)}$$

$$\text{Acc. rAAV in Lys (vg)} = \text{rAAV titer in Lys 96 hpt} \times V + \text{rAAV titer in SN 48 hpt} \times V \quad \textbf{(2)}$$

To determine the rAAV egress rate, the proportion of rAAV present in the supernatant fraction relative to the total rAAV was calculated using **Equation (3)**:

$$\text{Esn (\% SN)} = \frac{\text{rAAV titer in SN}}{\text{rAAV titer in Lys}} \times 100 \quad \textbf{(3)}$$

Cell cycle analysis

Freshly harvested cells were pelleted by centrifugation, rinsed, re-suspended in PBS and normalized to 1 x 10^6 cells in 500 µL. Cells were fixed with cold ethanol (70%) for 2 hours at -20ºC. For cell cycle analysis, cells were washed in PBS twice and stained with propidium iodide (PI) solution (50 µg/mL PI, 200 µg/mL RNAse A, 0.1% Triton X-100 (v/v)) for 30 minutes at 37ºC prior to data acquisition with CytoFLEX (Beckman Coulter Inc., Brea, CA, USA). Data was processed using FlowJo V10 (BD Biosciences, San Jose, CA, USA).

Protein samples preparation for mass spectrometry analyses

Proteomic analysis was performed at the Cardiovascular Proteomics group (CNIC, Madrid, Spain). Protein extraction was performed from pellet fraction using extraction buffer (100 mM Tris-HCl, 2 mM EDTA, 4% SDS, 100 mM DTT, at pH 8.8) of which 100 µL were added to the cell pellet of each condition. Samples were sonicated for 5 minutes and then boiled for another 5 minutes. Protein extracts from pellet samples were quantified with RC/DC Protein Assay (BioRad, Hercules, CA, USA) and stored at -20 ºC until the tryptic digestion process. Protein digestion was performed as previously described[302]. Proteins were digested using sequencing grade trypsin (Promega, Madison, WI, USA). The filter-assisted sample preparation technology (FASP, Expedeon, San Diego, CA, USA), and the resulting peptides were subjected to TMT-10 plex labelling (AB Sciex, Framingham, MA, USA), joined and desalted. A total of 150 µg from samples of each condition was diluted to a final concentration of 100 mM of TEAB labeled with TMT-10 plex according to the manufacturer. Protein samples were labeled by adding 41 µL of TMT isobaric tag diluted in anhydrous acetonitrile, followed by a 1 hour-incubation step at room temperature. To quench the reaction, 5 % hydroxyl-amine (8 µL per sample) was added, incubated 15 minutes at room temperature and mixed together followed by addition of TFA 1% to lower pH at 3. TMT-labeled samples were equally mixed. Pooled mix was purified using Oasis HLB C18 column (Waters, MA, USA). TMT-labeled mix was fractionated using high pH reversed-phase peptide fractionation kit

(ThermoFisher Scientific, San Jose, CA, USA) according to manufacturer's instructions into 5 fractions for further LC-MS/MS analysis.

Liquid chromatography tandem mass spectrometry analysis

The tryptic peptide mixtures were subjected to LC-MS/MS analysis on a nano-HPLC Easy nLC 1000 liquid chromatograph (ThermoFisher Scientific, San Jose, CA, USA) coupled to a QExactive mass spectrometer (ThermoFisher Scientific, San Jose, CA, USA). Peptides were suspended in 0.1% formic acid, loaded onto a C18 reverse-phase trapping column (Acclaim PepMap100, 75-µm internal diameter, 3-µm particle size and 2-cm length (ThermoFisher Scientific, San Jose, CA, USA), and separated on an analytical C18 nano-column (EASY-Spray column PepMap RSLC C18, 75-µm internal diameter, 3-mm particle size and 50-cm length (ThermoFisher Scientific, San Jose, CA, USA), in a continuous gradient (8–31%B in 240 min, 31–90%B in 2min, 90%B in 7 min, and 2%B in 30min; where buffer A is 0.1% formic acid in HPLC grade H_2O and buffer B is 100% ACN, 0.1% formic acid in HPLC grade H_2O). Spectra were acquired using full ion scan mode over the mass-to-charge (m/z) range 390–1500, 70,000 FT-resolution was performed on the top 15 ions in each full MS scan using the data-dependent acquisition mode with 45s dynamic exclusion enabled.

Protein identification and quantification

Protein identification was performed over the raw files using the SEQUEST HT algorithm integrated in the Proteome Discoverer 2.1 (Thermo Finnigan). MS/MS scans were matched against a human database (UniProtKB/Swiss-Prot 2019_10 Release).

For database searching, parameters were selected as follows: trypsin digestion with 2 maximum missed cleavage allowed, precursor mass tolerance of 800 ppm, fragment mass tolerance of 0.02 Da. TMT-10 plex labeling at N-terminal and lysine (+229.62932 Da) as well as cysteine carbamidomethylation (+57.021 Da) were chosen as static modifications whereas methionine oxidation (+15.994915 Da) was chosen as dynamic modification. The same MS/MS spectra collections were searched against inverted database constructed from the same target database. SEQUEST results were analyzed by the probability ratio method[303]. False discovery rate (FDR) for identified peptides was calculated in the inverted database search results using the refined method[304].

Quantitative information of TMT reporter ions was extracted from MS/MS spectra for relative quantification of protein abundance to characterize dynamic protein expression profiles in the selected conditions.

Statistical analysis

For the comparative analysis of the protein abundance changes Weighted Scan-Peptide-Protein (WSPP) statistical workflow[305,306], using SanXoT package was applied[307]. This method uses as an input a list of quantifications in the form of log2-ratios (for example a condition versus control sample) with their statistical weights and generates the standardized forms of the original variables computing the quantitative values expressed in units of standard deviation around the averages. The quantitative information is obtained from the spectra and used to quantify the peptides from which the spectra are produced and then, proteins that generate these peptides. In other words, the quantitative information is integrated from the spectrum level to the peptide level and then from the peptide level to the protein level[308]. These standardized variables, (Zq), express the quantitative values in units of standard deviation[305]. For the protein functional analysis, Systems Biology Triangle (SBT) model[308] was used. This algorithm estimates weighted functional category averages (Zc) from the protein values by performing the protein to category integration. After the integration from spectra to peptide and peptide to protein, this integration represents a higher level, from protein to category.

The integration allows the detection of changes in functional categories produced by the coordinated behavior of their proteins[308]. Together with each Zq and Zc, the corresponding FDR was calculated. 5% FDR was considered significant. The quantified proteins were functionally annotated using the Gene Ontology database[309,310]. For further Gene Ontology annotation, DAVID[311,312] was used to perform functional enrichment analysis. To help analyze and comprehend the data, the online software for reactions, proteins and pathways analysis REACTOME[313] was used. Additional bioinformatic analysis was performed using the STRING network interaction database (https://strin g-db.org/).

Western Blot

For Western Blot analysis, 600 µL of cell culture was centrifuged 300 x g for 5 minutes. Supernatant was discarded and cell pellet was resuspended in PBS and 1x cell lysate buffer for 1 hour at 37 ºC 130 rpm. Samples were centrifuged for 20 min at 4000 g at 4ºC and supernatant was then transferred into a clean tube on ice. Samples were treated with 1x Laemmli Sample Buffer (BioRad, Hercules, CA, USA), mixed and boiled for 20 minutes at 96ºC. Samples were ran on precast SDS-polyacrylamide (4–20%) gel electrophoresis at 200 V, 400 mA for 45 minutes. Precision Plus Protein Western C (#1610376, BioRad, Hercules, CA, USA) was used as molecular weight standard. Running buffer used was Tris/Glicine/SDS (25 mM Tris, 192 mM glycine, 0.1% SDS, pH

8.3) (BioRad, Hercules, CA, USA). Gel was transferred onto a polyvinylidene difluoride (PVDF) membrane for 7 minutes using the system Trans-Blot Turbo Transfer System (BioRad, Hercules, CA, USA). Membrane with transferred proteins were incubated for 1 hour in blocking buffer (1x PBS, 0.1% (v/v) Tween-20 with 5% (w/v) non-fat milk) followed by 1 hour incubation with the corresponding primary antibody at room temperature with gentle shaking. Primary antibodies used were Cap B1 (PROGEN Biotechnik GmbH, diluted 1:10), rabbit anti-CD63 antibody (ab134045, Abcam UK, 1:1000), mouse anti-CD81 antibody (ab79559, Abcam, UK, 1:1000), mouse anti-CD9 (ab223052, Abcam, UK, 1:1000) and rabbit anti-HMGCS1 (ab155787, Abcam, UK, 1:200). Membrane was washed three times for 5 minutes on orbital shaker with wash buffer (1x PBS, 0.1% (v/v) Tween-20). After primary incubation, a secondary incubation was performed using HRP-linked anti-mouse or anti-rabbit antibody accordingly (A3562, Sigma-Aldrich, St. Louis, MO, USA) at 1:2500 in 2.5% (w/v) nonfat dry milk 1x PBS 0.1% (v/v) Tween-20 for 1 hour at room temperature. Proteins were visualized using Clarity Western ECL Substrate solution (#1705061, BioRad, Hercules, CA, USA) using ChemiDoc MP Imaging System (BioRad, Hercules, CA, USA).

Experimental design and statistical rationale

For the multiplexed quantitative proteomics experiment based on TMT-10 plex labeling, three conditions were tested, with three biological replicates (independent cell cultures) each. The experiment included a non-transfected group maintained as control and two rAAV9 production conditions: standard TT protocol for batch production and extended gene expression (EGE) protocol, characterized by a medium exchange (ME) and a subsequent re-transfection step (RT) at 48 hpt as illustrated in **Figure 1A**. Samples from each condition were taken at three different time points: 48, 72 and 96 hpt. To note, sample at 48 hpt was taken before the EGE protocol steps of ME and RT. The TMT-based isobaric labelling quantification was performed with three biological replicates for each condition and time point. All individual values were normalized by the value from a peptide pool from the non-transfected condition at 48hpt which was used as an internal control for the all the statistical analyses.

RESULTS

The extended gene expression protocol enhances rAAV production compared to standard triple transfection

Cell cultures were inoculated at 0.5 x 10^6 cells/mL and transfected when reaching 2 x 10^6 cells/mL. The cell growth condition, which was not transfected, was maintained as a control. Two different rAAV9 production conditions were compared: the standard triple transfection (TT) protocol for batch production and an enhanced rAAV9 production protocol based on extended gene expression (EGE), characterized by a medium exchange (ME) and re-transfection (RT) step at 48 hpt (**Figure 1A**). As shown in **Figure 1B**, the rAAV9 production conditions reached a cell concentration of 4 x 10^6 cells/mL. In contrast, the non-transfected condition exhibited continued growth beyond the mid-exponential phase, reaching a high cell density of 6 x 10^6 cells/mL[314]. The EGE protocol sustained a cell viability of 80%, whereas the TT protocol showed a decrease to 60% (**Figure 1B**).

The accumulated rAAV9 production in cell supernatant and cell lysate fractions was analyzed as described in **Equation 1** and **2**. Following the EGE protocol, a marked increase in total accumulated titer of rAAV9 compared to TT protocol was observed in supernatant and lysate harvest fractions (**Figure 1C**). Additionally, the proportion of rAAV9 harvested into the supernatant was 49% with the EGE protocol, significantly higher than the 35% observed in the TT production (**Figure 1D**). This comparison indicates that the EGE protocol is associated not only with increased production titers but with enhanced efficiency of rAAV9 viral egress.

Proteomic landscape upon rAAV production protocols

A comparative analysis of altered cellular processes during cell growth, standard TT rAAV production, and the EGE protocol was conducted using a multiplexed quantitative proteomic approach (schematically defined in **Figure 1A**). Individual proteins were quantified showing differential protein changes across each studied condition at 48, 72 and 96 hpt. A total of 9,255 proteins were identified and quantified. From these, 6,497 proteins were quantified with more than 1 peptide. 81.5%, 87.4% and 83.4% of proteins from non-transfected, TT and EGE conditions respectively were present in all three conditions and quantified with more than 1 peptide.

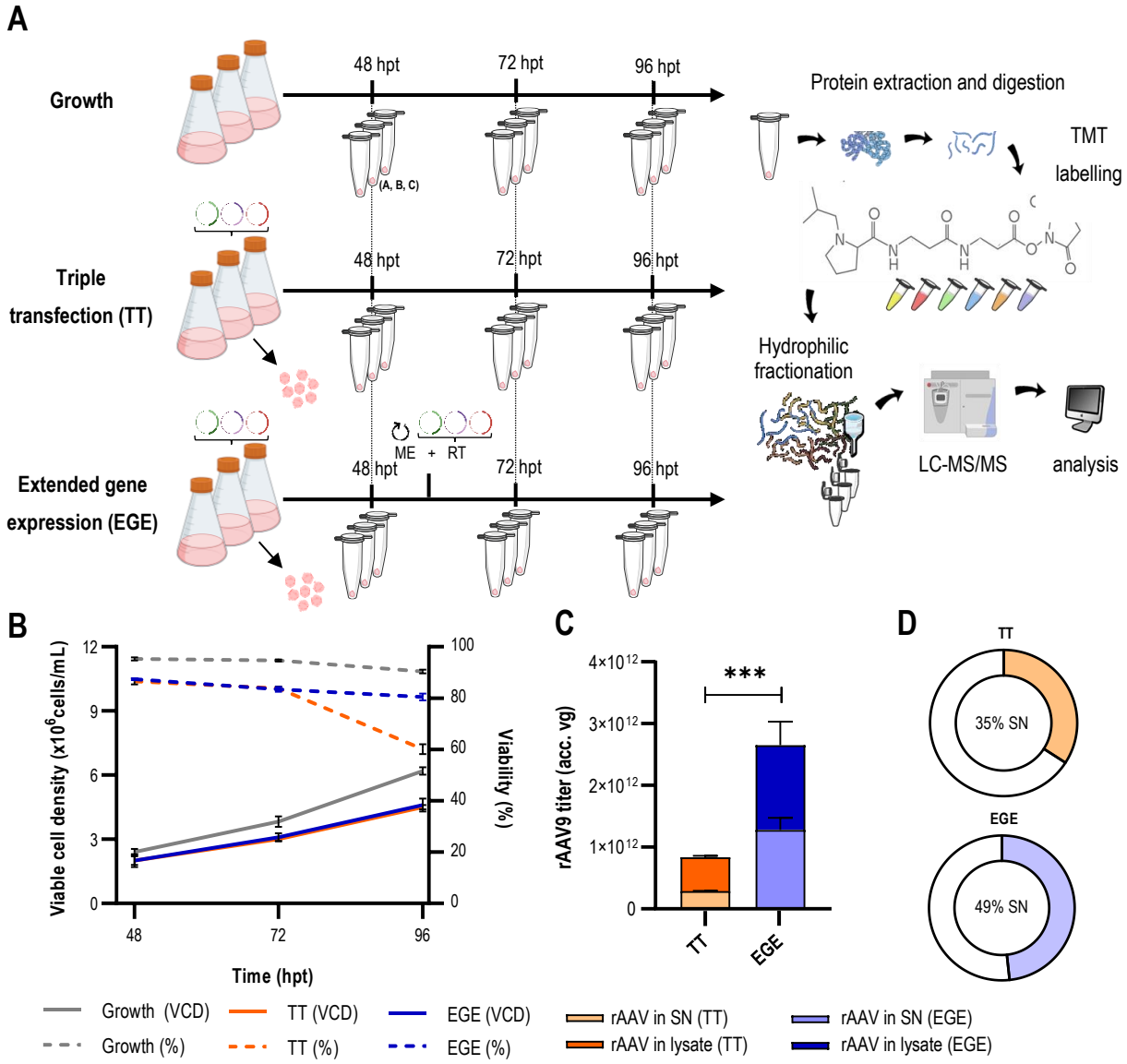


Figure 1. Experimental design and rAAV9 production analysis. A. Experimental workflow. Three biological replicates (A, B and C) of a HEK293 cell growth condition and two rAAV9 production conditions: standard triple transfection (TT) protocol and extended gene expression (EGE) protocol, characterized by a medium exchange (ME) and re-transfection step at 48 hours post-transfection (hpt). Cell culture samples were taken at 48, 72 and 96 hpt. Proteins were extracted from the pellet samples, digested and peptides were labelled using Tandem Mass Tag (TMT) labelling. Labelled peptides were fractionated and analyzed via LC-MS/MS. **B.** Viable cell density and viability analyzed during production time course. Cells were transfected at $2x10^6$ cells/mL at 0 hpt. **C.** rAAV accumulated in cell lysate and cell supernatant 96 hpt from a 10 mL cell culture. **D.** Fraction of total rAAV9 harvested 96 hpt accumulated in cell supernatant fraction (SN). Each different condition in triplicate was measured in triplicate (n=3). Data was analyzed by one-way analysis of variance (ANOVA) for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. EGE, extended gene expression. ME, medium exchange. RT, re-transfection. SN, supernatant. TT, triple transfection. VCD, viable cell density. vg, vector genomes.

This accounted for a total of 4679 proteins (**Figure 2A-B**). All analyses, except if specifically mentioned, were carried out using this set of 4679 common proteins. Some proteins were exclusively identified in two conditions such as the AAV proteins Rep78/68 and capsid proteins, validated by Western Blot (**Figure 2C**). Adenoviral proteins E1A and E1B, stably integrated in HEK293 cell genome were identified in the three experimental conditions and showed significant differential changes. A more detailed analysis on E1 adenoviral protein expression is presented in **Chapter three**.

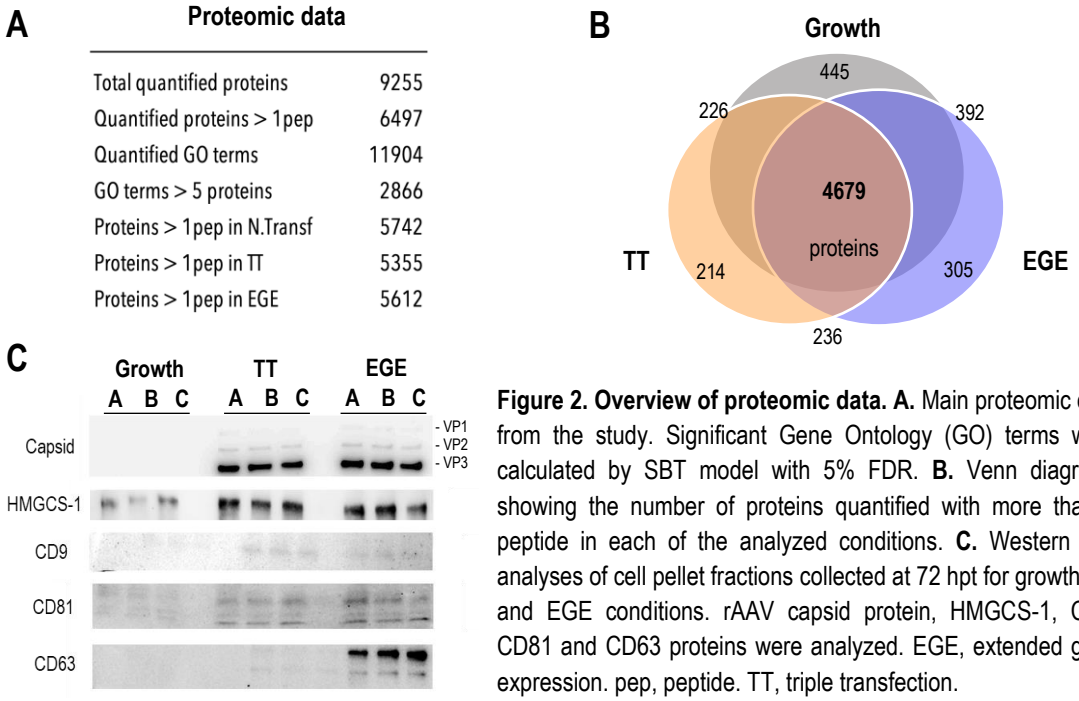


Figure 2. Overview of proteomic data. A. Main proteomic data from the study. Significant Gene Ontology (GO) terms were calculated by SBT model with 5% FDR. **B.** Venn diagrams showing the number of proteins quantified with more than 1 peptide in each of the analyzed conditions. **C.** Western Blot analyses of cell pellet fractions collected at 72 hpt for growth, TT and EGE conditions. rAAV capsid protein, HMGCS-1, CD9, CD81 and CD63 proteins were analyzed. EGE, extended gene expression. pep, peptide. TT, triple transfection.

Here, a detailed analysis focusing on individual protein and biological changes comparing TT and EGE protocols with control growth condition is presented. Up and downregulations affecting cell growth condition from 48 to 96 hours of cell culture were attributed to cell growth effect, providing insight of metabolic alterations in HEK293 cultures under proliferation. Changes in both rAAV production conditions 48 hpt in comparison to growth condition were explained by rAAV production process, as well as alterations over time during transient transfection method. Finally, changes in EGE condition in comparison to TT at 72 and 96 hpt were attributed to rAAV9 production methodology, concretely to ME and RT step performed at 48 hpt. Thus, shedding light in processes involved in enhanced rAAV9 production. The following analysis were carried out following these guidelines.

Gene expression processes: transcription, RNA processing, and protein production homeostasis regulation over rAAV production

The journey from a gene to a functional protein in a cell is a multi-step process, each of them governed by different regulation mechanisms. The dynamics of transcription, RNA processing, translation and protein post-translational modification are key to finally obtain a fully functional protein. Over multiple plasmid transfection for rAAV production, these processes are exploited by the viral genes to commit the host cell to produce their own proteins. In this context, transcription is a key process that affects the first steps of gene expression.

The obtained results showed a decrease in cellular transcription processes upon rAAV9 production, as depicted in **Figure 3A** at 72 hpt compared to the cell growth condition. This downregulation was accentuated in the EGE protocol, with a marked decrease in RNA polymerase II regulation, mRNA splicing and transcription factor activity. In contrast, viral transcription increased in rAAV production protocols (**Figure 3B**). This suggests a shift in the cellular machinery focus from maintaining normal cellular functions, as in cell growth, towards facilitating viral production.

Molecular events affecting RNA processing and protein translation and homeostasis response were also studied (**Figure 3B, C**). During cell proliferation, both RNA processing and protein translation and homeostasis responses decreased. However, these processes tended to be upregulated during rAAV production. Protein homeostasis was notably altered; in particular, ATF6 and PERK, key components of the unfolded protein response (UPR) pathways were upregulated upon rAAV production in comparison to cell growth condition. Other processes involved in protein folding in the endoplasmic reticulum (ER) as well as chaperone-mediated folding, were also upregulated during the TT rAAV production, inverting the trend observed in cell growth (**Figure 3B, C**).

Furthermore, significant alterations in proteins associated with molecular processes related to gene expression were observed between the two rAAV production methods. To this end, a comparison of protein expression between the EGE protocol and TT at 72 hpt was conducted (**Figure 3D**). A downregulation of the transcription cofactor MAF1 was observed in EGE compared to TT, evidencing the downregulation in transcription processes. Conversely, proteins involved in mRNA translation initiation and elongation, such as EIF2B and EEF1A1, were upregulated in the EGE protocol (**Figure 3D**). Furthermore, the EGE demonstrated significant alterations in proteins related to the cellular homeostasis response. Remarkably, cytoplasmic chaperones such as HSPA1B were upregulated in the EGE protocol, whereas chaperones involved in protein folding in the ER, such as calreticulin (CALR), were significantly downregulated (**Figure 3D**). Another significant

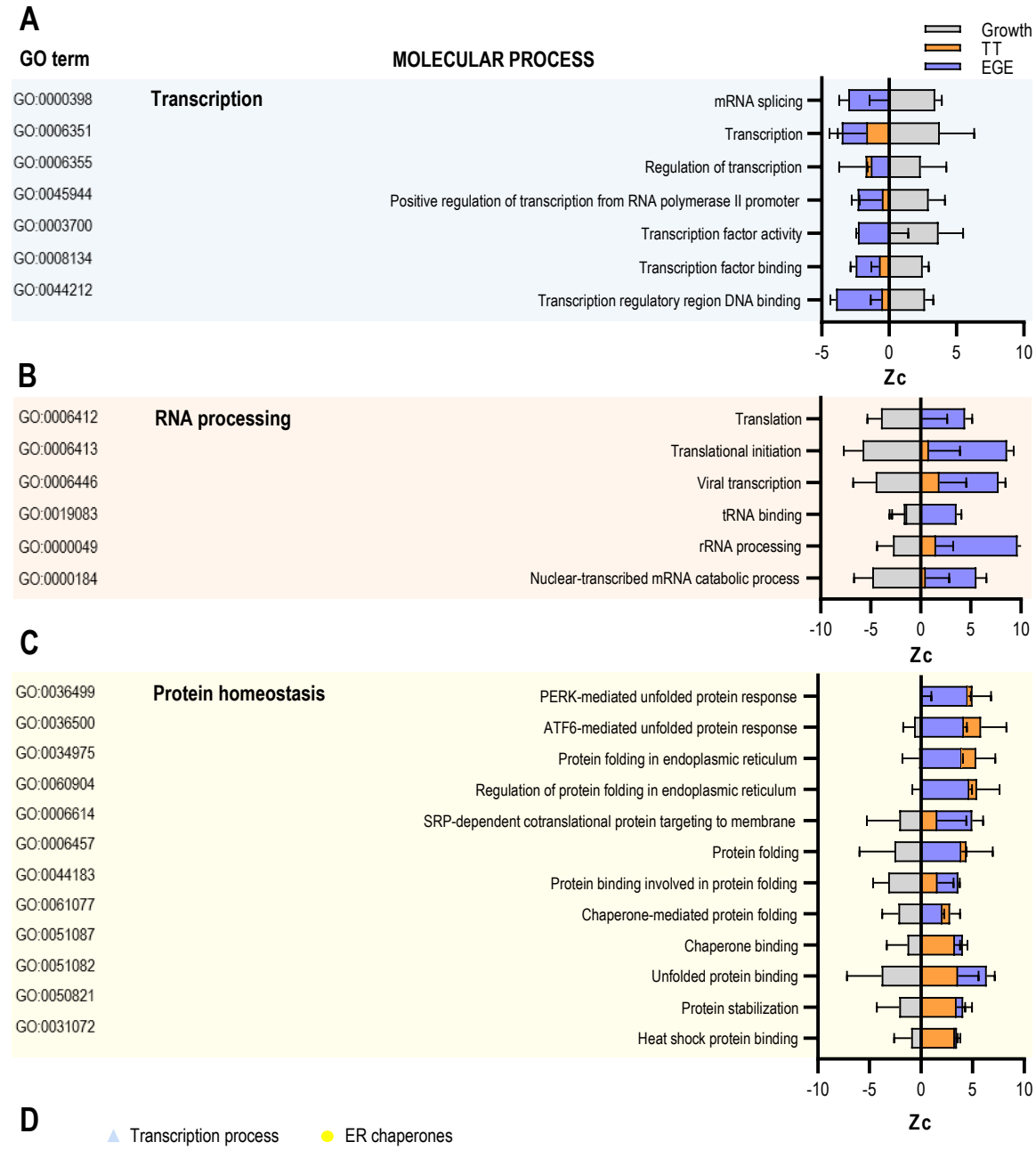


Figure 3. Altered biological processes related to transcription, RNA processing and protein production and homeostasis. Bar plots representation of significant altered processes (Zc values) due to rAAV production. Comparison between growth and TT condition 72 hpt. EGE condition was also graphed. Significance was calculated using two-way ANOVA using Sidak's multiple comparison test. **A.** Processes related to transcription. **B.** Processes related to mRNA translation. **C.** Processes related to protein production and homeostasis. **D.** Volcano plot representing altered proteins of EGE condition in comparison to TT condition 72hpt. Different colors and shapes in figure legend correspond to the different analyzed group of proteins. EGE, extended gene expression. TT, triple transfection.

observation was the notable upregulation of proteins associated with protein tagging for degradation, such as ubiquitin-ligases (UBE2C), in the EGE protocol compared to TT. The Sequesterome (SQTM1) protein, involved in autophagy, was also upregulated. Additionally, proteins related to cell homeostasis and cell cycle progression, including Tyrosin-kinase (JAK1), related to immune processes, and retinoblastoma protein (RB1), involved in G1/S cell cycle progression, were increased in the EGE protocol (**Figure 3D**). Overall, these differential changes in EGE compared to TT, underscore the extensive cellular reprogramming regarding the gene expression-related processes and cell homeostasis regulation pathways under the EGE protocol, potentially contributing to its enhanced rAAV production capabilities.

Altered lipid and glycosphingolipid metabolic processes in rAAV production protocols compared to cell growth

Lipid biosynthetic processes showed a downregulation during cell growth. However, an upregulation in lipid biosynthetic pathways was observed during rAAV production (**Figure 4A-D**). Although the TT protocol initially showed a reduction in lipid biosynthesis (72 hpt), levels were restored at 96 hpt. Contrarily, the EGE method consistently upregulated lipid biosynthetic pathways over time (**Figure 4A**). These alterations in terms of specific lipid-related molecular processes are depicted in **Figure 4D**. Specifically, the biosynthesis of fatty acids, lipids and cholesterol showed a similar upward trend while glycosphingolipid metabolism remained unaffected upon different rAAV production protocols.

Subsequent analysis focused on the modulation of proteins by comparing the EGE and TT protocols at 72 hpt (**Figure 4B, C**). In this context, expression levels of proteins within lipid biosynthetic pathways were downregulated during cell growth, and remarkably upregulated in the EGE condition compared to TT. This included key enzymes such as fatty acid synthase (FASN), and components of the cholesterol synthetic pathway, including squalene synthase (FDFT1) and hydroxymethylglutaryl-CoA synthase (HMGCS1). Conversely, in the TT protocol, these proteins showed a minor increase, restoring their levels to baseline (**Figure 4B**).

Interestingly, when examining proteins related to the glycosphingolipid metabolic process, the EGE protocol showed a differential regulation pattern compared to the TT approach, varying by the specific protein involved (**Figure 4C**). Generally, glycosphingolipids were found to be downregulated during rAAV production. However, certain enzymes such as sphingomyelin phosphodiesterase 4 (SMPD4) and glycolipid transfer protein (GLTP) exhibited an upregulation in the EGE method, increasing the levels observed in standard rAAV production.

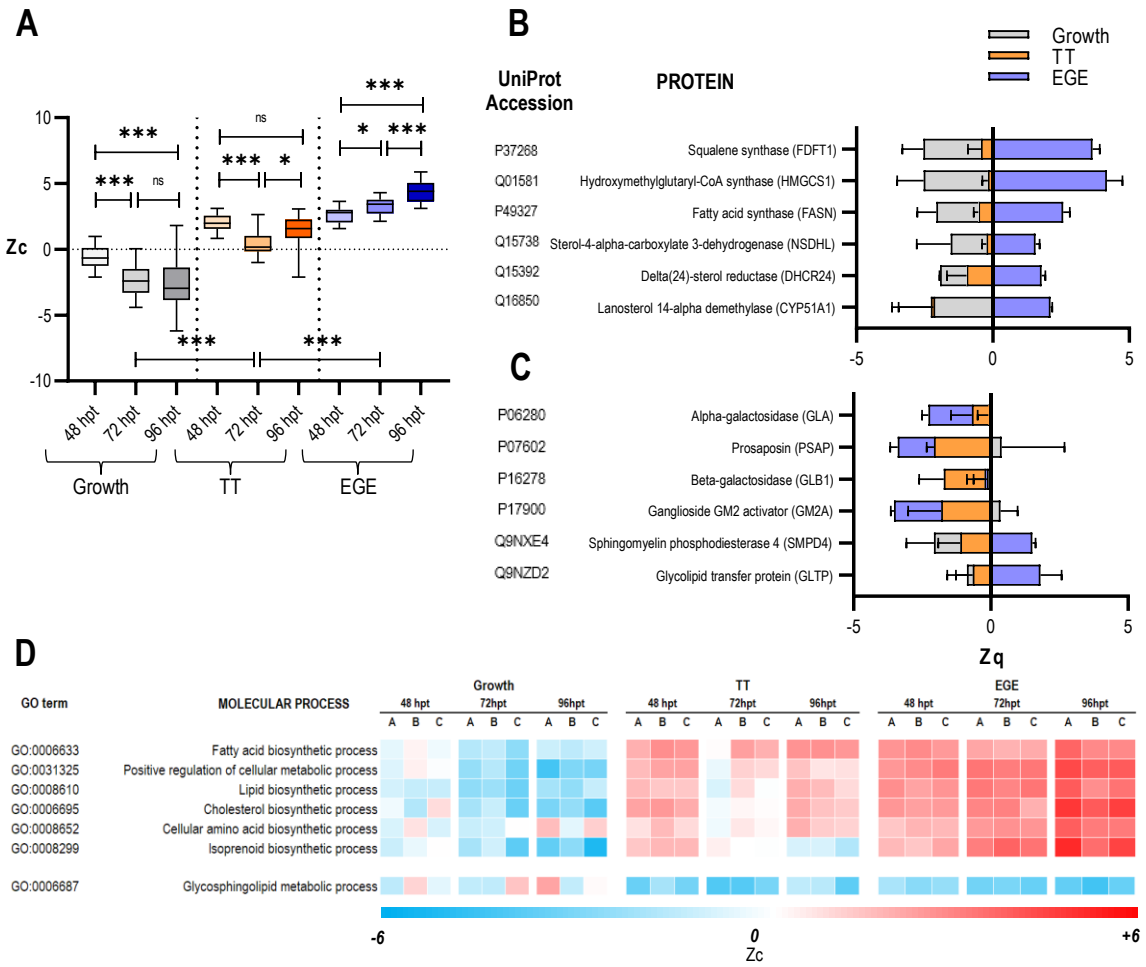


Figure 4. Altered biological processes and proteins related to lipid and glycosphingolipid metabolism. A. Box plots representations of the Zc values from GO terms related to lipid biosynthesis. Comparison of different time points (48, 72, 96hpt) in the three different conditions (growth, TT and EGE). Medians are represented by vertical bars and whiskers extend to extreme data points. Significance was calculated using one-way ANOVA and Tukey's comparison test. **B.** Bar plots representation of significant altered proteins (Zq) upregulated in EGE in comparison to TT condition 72hpt. Growth condition is also illustrated. **C.** Bar plots representation of significant altered proteins related to glycosphingolipid pathway (Zq) hindered in EGE compared to TT condition 72hpt. Growth condition is also illustrated. Significance was calculated using two-way ANOVA using Sidak's multiple comparison test. **D.** Heat map of Gene Ontology (GO) biological processes terms regarding lipid biosynthesis and quantified with more than 5 proteins. Zc values are represented following the color-coding legend found below, red for upregulation and blue for downregulation. A, B, C, represent the three biological replicates. EGE, extended gene expression. TT, triple transfection.

Analysis of HEK293 energy-related processes

Mitochondrial protein analysis revealed notable changes in cell's energy demands and metabolic preferences across different conditions, including cell growth and distinct rAAV production methods. During cell growth and the TT condition, an upregulation of mitochondrial proteins was observed, indicating increased metabolic activity and energy requirements. In contrast, a significant

decrease in mitochondrial protein expression was noted during the EGE protocol, particularly at 72 hpt (**Figure 5A**).

Further analysis comparing the EGE condition to the TT method at 72 hpt, revealed a marked downregulation of mitochondrial proteins in the EGE method. These proteins can be categorized based on their functions, with major changes associated with energy production and the electron transport chain (ETC) proteins. Key ETC proteins, such as cytochrome c1 (CY1) and ATP synthase subunit d (ATP5H) displayed reduced expression in the EGE condition relative to the TT (**Figure 5B**). This suggests a potential shift in energy metabolism or mitochondrial dynamics triggered by the EGE's differential protocol steps.

Intracellular energy transfer is also affected by importin proteins, which play a pivotal role as regulators of intracellular transport. Distinct expression patterns of cellular and mitochondrial importins were observed (**Figure 5C**). Cellular importins primarily mediate the transfer of molecules across the nuclear envelope, facilitating communication and material exchange between the nucleus and the cytoplasm. On the other hand, mitochondrial importins specialize in the transport of proteins from the cytosol into the mitochondria, playing a vital role in maintaining mitochondrial function[315].

A detailed analysis of all altered importin proteins was conducted at 72 hpt across the three studied conditions (**Figure 5D**). Cellular importins were downregulated over time during cell growth. However, significant differences in expression were observed between the different rAAV production protocols. While the TT method did not alter the expression of cellular importins, the EGE method upregulated them in comparison to the TT condition. Specifically, the importin subunit α-1 (KPNA2) was highly upregulated upon the EGE protocol.

Additionally, distinct expression patterns for mitochondrial importins were identified and classified into two groups (I and II), as shown in **Figure 5D**. Mitochondrial importins from group I exhibited a significant decrease in expression under the EGE protocol, particularly, TIMM8B, 8A, and TOMM70 showed marked reductions. In contrast, mitochondrial importins from group II did not show notable expression changes across the three studied conditions at 72 hpt.

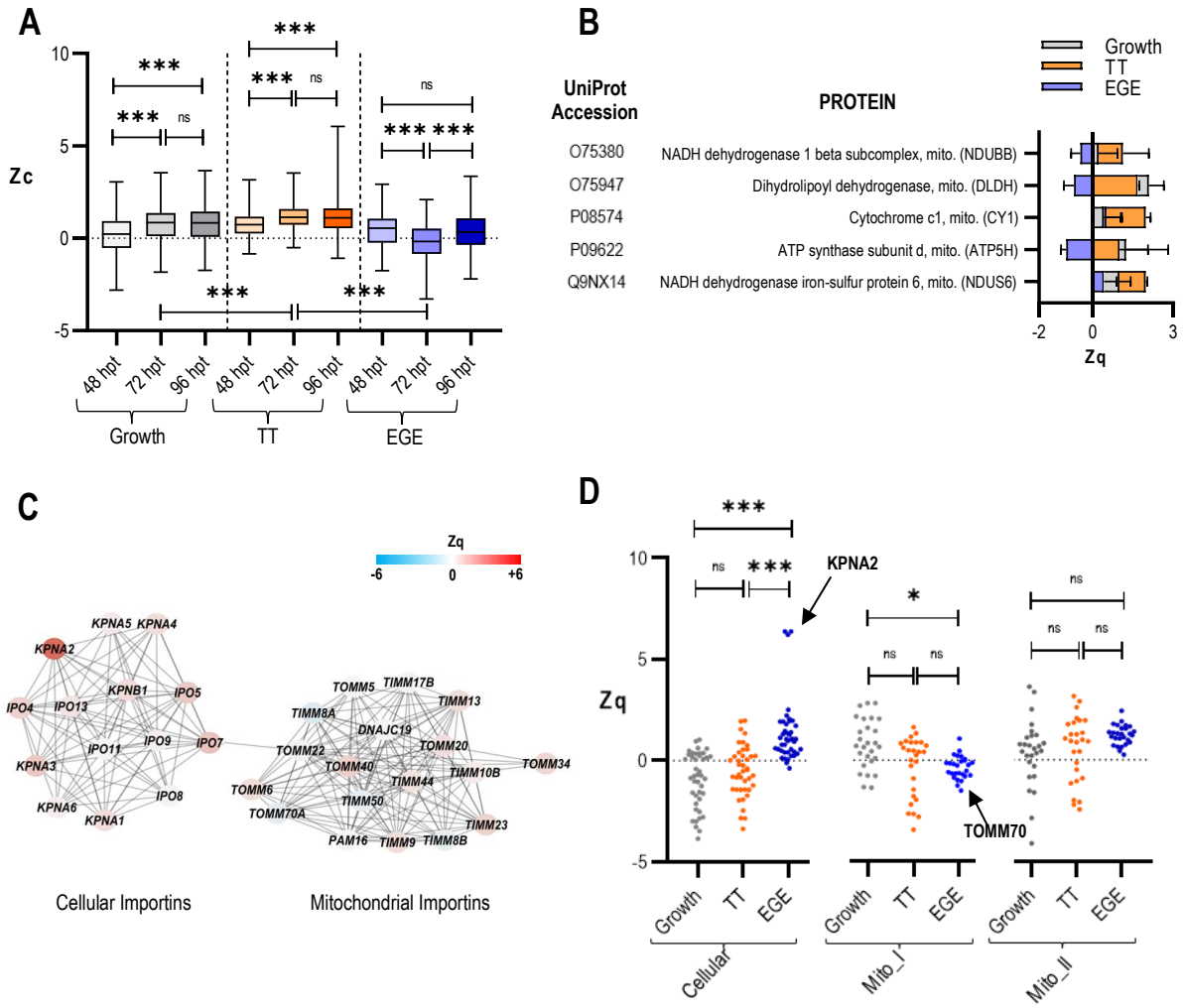


Figure 5. Altered biological processes and proteins regarding mitochondrial content and cellular and mitochondrial importins. A. Box plots representations of the Zc values from GO terms related to mitochondrial proteins. Comparison of different time points (48, 72, 96hpt) in the three different conditions (growth, TT and EGE). Medians are represented by vertical bars and whiskers extend to extreme data points. Significance was calculated using one-way ANOVA and Tukey's comparison test. **B.** Bar plots representation of significant altered mitochondrial proteins related to electron transport chain (ETC) (Zq) upregulated in EGE in comparison to TT condition 72hpt. Significance was calculated using two-way ANOVA using Sidak's multiple comparison test. **C.** String clustering analysis of cellular importins and mitochondrial importins showing variation upon different rAAV production processes. Zq values are represented following the color-coding legend, red for upregulation and blue for downregulation. **D.** Dot plot representation of cellular and mitochondrial proteins (group I and II) of Growth, TT and EGE conditions compared 72hpt. Significance was calculated using one-way ANOVA and Tukey's comparison test. EGE, extended gene expression. Mito, mitochondrial. TT, triple transfection.

Analysis of HEK293 cell cycle and DNA replication during rAAV production

Through detailed examination of the identified proteins and their significance within cell cycle checkpoints, the effects of rAAV production on the cell cycle progression and DNA replication during growth and rAAV production processes were studied (**Figure 6A-D**).

Analysis of all biological processes related to cell cycle regulation showed no differences over time in cell growth condition or in the TT rAAV production condition. Contrarily, a slight increase in cell cycle processes was observed during the EGE protocol compared to TT at 72 hpt, with a broad variability range (**Figure 6A**). To further investigate these alterations, a deeper study regarding the different cell cycle phases in EGE compared to TT was performed at 72 hpt (**Figure 6C**).

Molecular processes related to the G1/S phase transition of the cell cycle showed the most significant alterations. The TT method displayed a pronounced decrease in the G1/S transition, which was prevented by the EGE protocol. Consistently, the EGE protocol showed a decrease in negative regulators of the G1/S phase, thereby reducing the mechanisms that would inhibit cell cycle progression (**Figure 6C**). Overall, differential regulation of G1/S phase transition suggest that the EGE protocol counteracts the downregulation of the S phase transition induced by the rAAV production in the TT setting.

In addition, the EGE protocol markedly upregulated the G2/M phase transition, indicating a potential for rapid cell cycle progression. The regulation of mitotic cell cycle was also upregulated in EGE condition. Further, processes related to ubiquitin-protein ligase activity, involved in mitotic cell cycle, which play a crucial role in tagging proteins for degradation and enabling timely regulated cell cycle progression, showed a significant increase in EGE and a marked decrease in TT condition (**Figure 6C**). This implies an enhanced rate of protein turnover under the EGE protocol.

As observed in other studies and validated here, cells in culture achieving high cell densities, upregulate DNA repair and replication proteins, as observed in **Figure 6B**[316]. With TT rAAV production, there is a downregulation of DNA replication machinery. Differences between the two rAAV production methods were noted: while DNA replication is slightly upregulated over time in TT condition, the EGE production method tends to decrease DNA replication processes.

To further characterize cell cycle dynamics in the three studied conditions, flow cytometry analysis was performed 72 hpt (**Figure 6D**). In rAAV production, both TT and EGE showed a strong deregulation of the cell cycle, characterized by a wider peak in G1 phase compared to cell growth, suggesting a delay or arrest in this cell cycle phase. Moreover, S phase in rAAV production methods was narrower than cell growth condition, indicating a reduced progression of cells into the DNA synthesis phase during rAAV production, consistent with proteomic data.

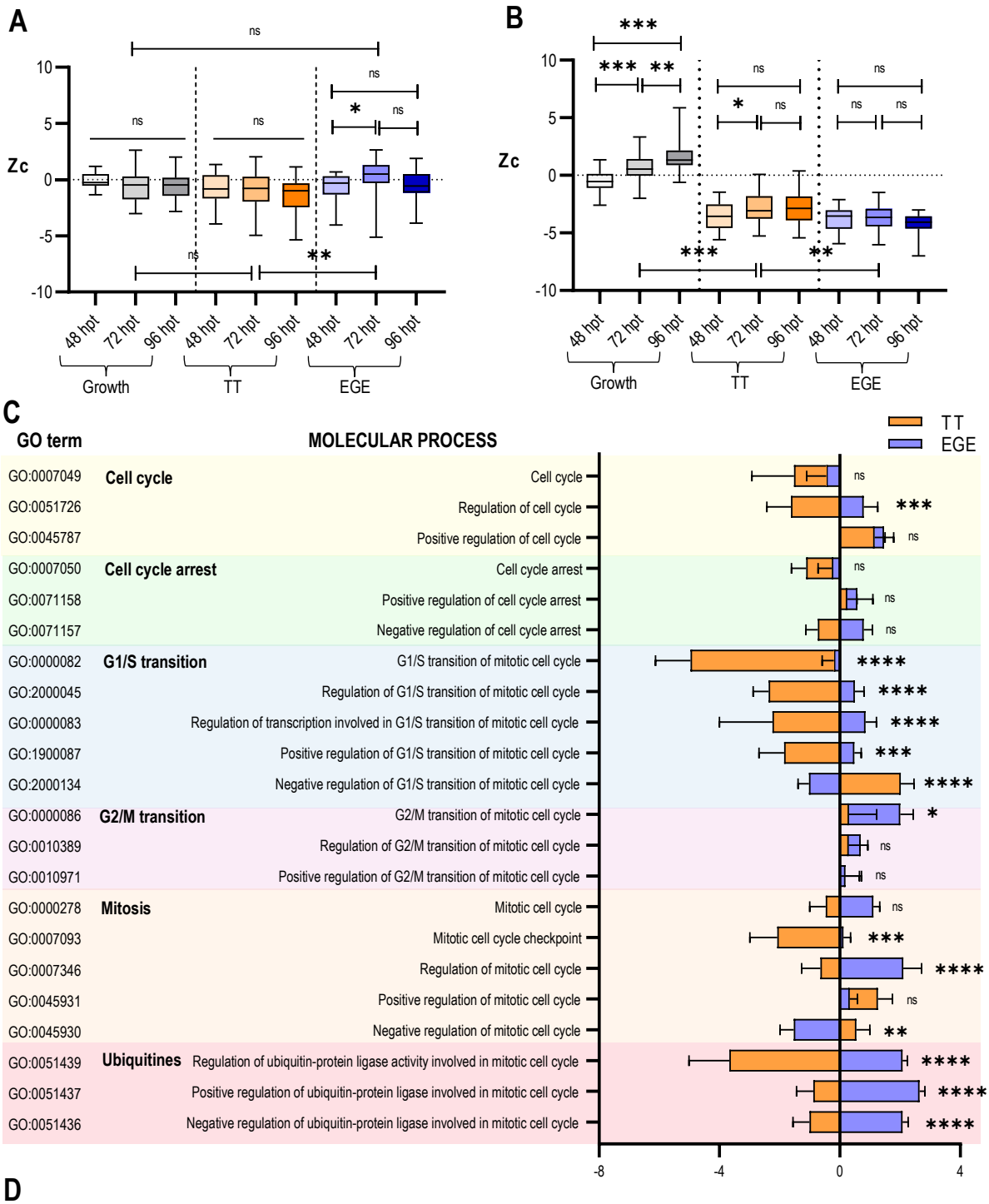


Figure 6. Analysis of cell cycle regulation and DNA replication. A. Box plots representations of the Zc values from GO terms related to cell cycle regulation. **B.** Box plots representations of the Zc values from GO terms related to DNA replication. Comparison of different time points (48, 72, 96hpt) in the three different conditions (growth, TT and EGE). Medians are represented by vertical bars and whiskers extend to extreme data points. Significance was calculated using one-way ANOVA and Tukey's comparison test. **C.** Bar plot representation of processes related to cell cycle regulation of EGE condition in comparison to TT condition 72hpt. **D.** Effect of Growth, TT and EGE conditions on cell cycle. Histogram representation of cell cycle phases G1, S and G2. EGE, extended gene expression. TT, triple transfection.

Analysis of HEK293 vesicle transport pathways

Vesicle transport pathways play a crucial role in maintaining cellular homeostasis and responding to environmental changes, particularly by regulating the intracellular trafficking of molecules between different compartments using vesicles. Vesicles are small, membrane-bound sacs that can bud off from one membrane and fuse with another, enabling the transport of their contents within the cell.

Under cell growth conditions at 96 hpt, a decrease in molecular processes related to vesicle transport pathways was observed. However, during rAAV production, an initial upregulation in vesicle transport pathways was noted, which varied over time and depending on the rAAV production protocol. The TT method showed a downregulation, while the EGE approach remarkably enhanced vesicle transport pathways over time, especially at 96 hpt (**Figure 7A**).

Subsequent analysis of molecular processes related to vesicle transport pathways revealed upregulated processes due to rAAV production. Specifically, pathways associated with ER to Golgi and intra-Golgi vesicle-mediated transport pathways were markedly enriched. In contrast, COPII vesicle coating did not show relevant differences compared to growth condition. Processes related to Cajal body regulation exhibited an upregulation (**Figure 7B**). Noteworthy, the EGE method showed a marked increase in certain molecular processes compared TT protocol. After ME and RT steps, intra-Golgi mediated transport, vesicle-mediated transport and membrane fusion was increased. The formation and binding of microtubules were also elevated.

Overall, these results suggest an enhanced vesicle transport pathways in the EGE protocol compared to TT. Consequently, an in-depth analysis of proteins involved in secretory vesicle transport pathways was conducted (**Figure 7C**). The endosomal sorting complex required for transport (ESCRT) machinery was identified and quantified. The ESCRT machinery is critical in the formation of intraluminal vesicles (ILV) within multivesicular bodies (MVB) post the trans-Golgi network stage. Here, the ESCRT-I protein VPS37B showed a significant increase in the EGE condition compared to TT at 72 hpt. Other mechanisms for MVB formation, ESCRT-independent such as tetraspanins and lipid-mediated MVB formation were also studied[317]. For transmembrane tetraspanins, proteins such as CD63, showed upregulation in the EGE method (**Figure 7C**). Further, molecular machineries involved in exosome or extracellular vesicle secretion were analyzed based on proteomic results upon rAAV production using the two studied methods[317]. Ras-related proteins named Rab GTPases, which are key regulators of intracellular vesicle transport, and involved in vesicle budding, mobility through the cytoskeleton of tethering to the acceptor membrane component were upregulated 72 and 96 hpt in the EGE protocol (**Figure 7C**). Finally,

SNARE proteins (for soluble N-ethylmaleimide-sensitive fusion attachment protein (SNAP) receptors) form complexes with SNAPs between two membranes and mediate membrane fusion between the two organelles. In this study, SNAP47 protein increased upon the application of the EGE protocol compared to TT (**Figure 7C**).

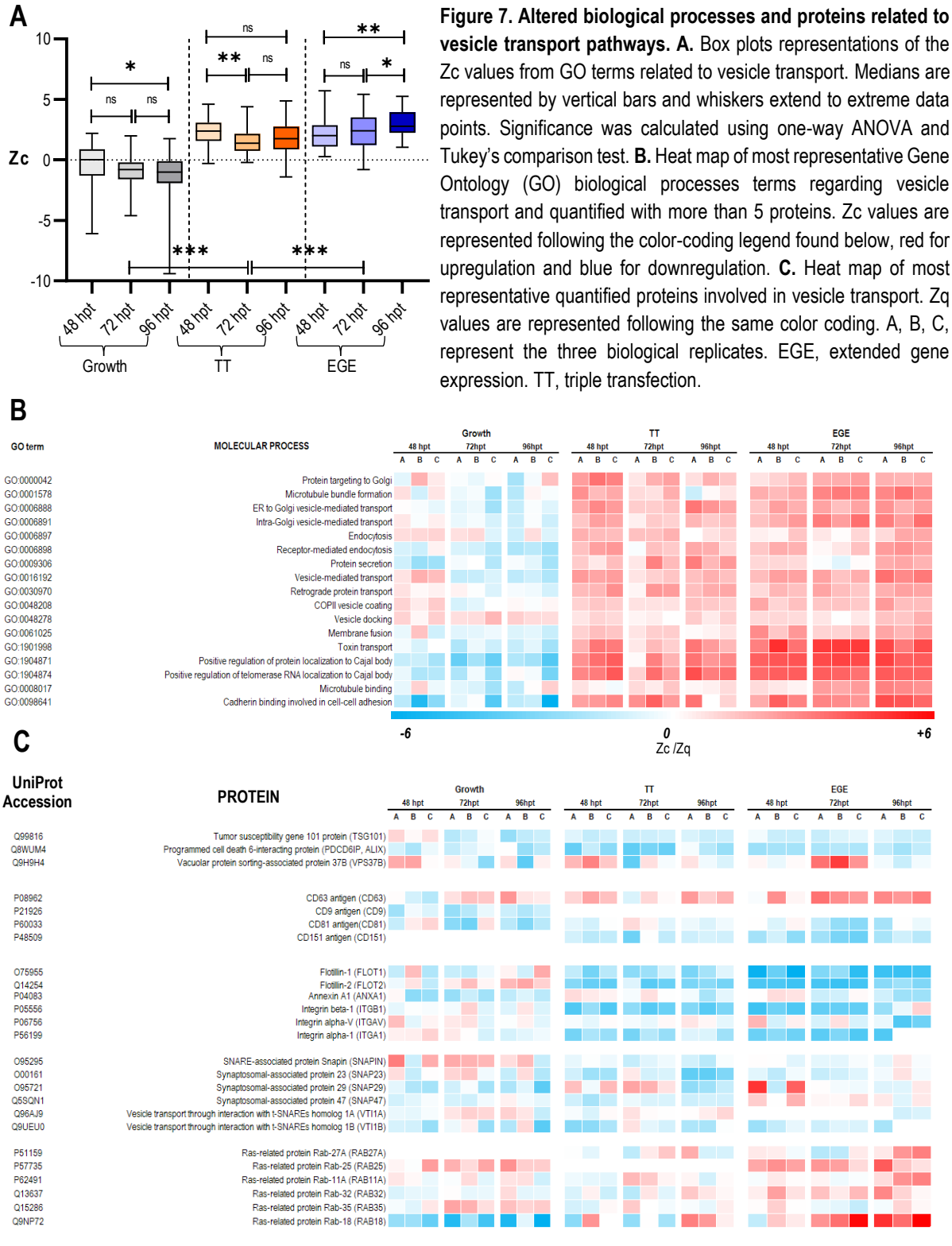


Figure 7. Altered biological processes and proteins related to vesicle transport pathways. A. Box plots representations of the Zc values from GO terms related to vesicle transport. Medians are represented by vertical bars and whiskers extend to extreme data points. Significance was calculated using one-way ANOVA and Tukey's comparison test. **B.** Heat map of most representative Gene Ontology (GO) biological processes terms regarding vesicle transport and quantified with more than 5 proteins. Zc values are represented following the color-coding legend found below, red for upregulation and blue for downregulation. **C.** Heat map of most representative quantified proteins involved in vesicle transport. Zq values are represented following the same color coding. A, B, C, represent the three biological replicates. EGE, extended gene expression. TT, triple transfection.

DISCUSSION

In this work, the effects of rAAV production on HEK293 cells are characterized using a multiplexed quantitative proteomic approach. The experimental setup included a non-transfected cell growth condition and two rAAV9 production protocols: the standard TT and the EGE at three time points: 48, 72 and 96 hpt. This design allows the interpretation of protein changes and the associated biological effects grouped by Gene Ontology (GO) terms summarized in **Figure 8A**.

Overall, the proteomic differences identified in this study are discussed in its biological context to provide insights into host cell modulations. This analysis aims to inform future rational engineering of host cells to enhance rAAV production by targeting specific steps in the process. Proteomic changes during cell growth condition are attributed to the CDE understanding, recognized as an important bottleneck affecting the TGE scale-up. Differential protein changes in rAAV production by the standard TT protocol compared to the cell growth condition, highlight cellular reprogramming induced by rAAV production. Specific alterations occurring post-ME and RT steps, at 72 and 96 hpt in the EGE method, underscore the differential capacity of this protocol to modulate host cell biological effects. These changes are discussed in two contexts. First, the host cell response to the EGE method may be attributed to its enhanced capacity to increase rAAV production compared to TT. Moreover, the ME step in the EGE method compared to cell growth condition, may be of use considering its contribution modulating the CDE. A summary of the hypothesized mechanisms by which the EGE outperforms TT at a molecular level is detailed in **Figure 8B** and discussed below.

Modulation of the cell density effect in HEK293 cells

Previous studies have explored the relationship between the CDE and energy metabolism, reporting that ATP production rates decrease at high cell densities[281,298,318]. In contrast, perfusion systems which allow for continuous media exchange are related with a high ATP production rates, which has been linked to help attenuate the CDE to achieve successful infections at HCD[285].

In this study, media exchange was performed once by centrifugation. Collecting data at discrete time points allowed for the characterization of cellular responses over the ME step. Thus, the analysis of the cell growth condition from 48 to 96 hpt, where HCD is observed, enables to elucidate molecular pathways associated with the CDE. Similarly, by performing ME 48 hpt in the EGE method of rAAV production allowed changes at 72 and 96 hpt to be correlated with mechanisms associated with the host cell response to overcome the CDE.

Consistent with previous reports, mitochondrial importins and mitochondrial proteins were upregulated at HCD, reflecting an increase in energy demand, leading to higher protein import in the mitochondria[316]. However, in the EGE method, mitochondrial proteins were downregulated after the ME and RT step. This differentially regulated metabolism might suggest that the cells are reducing their mitochondrial protein content. This is further validated by the observation that mitochondrial importins are as well downregulated after ME and RT steps. This could be indicative of a diminished need for protein import into the mitochondria that might be related to a decreased energetic requirement. However, the specific analysis of energy metabolism could generate insights into ATP production rates in each rAAV production state.

Similarly, lipid biosynthetic pathways downregulation observed upon HCD in cell growth condition could be linked to energy conservation strategies, as de novo lipid synthesis is a high-energy consuming process. This was also observed by Lavado *et al*. in similar growth conditions[316]. Cell membrane lipid composition might be related to the capability of the cell to internalize DNA/PEI complexes[316]. Importantly, the implementation of the EGE method contributed to counteract the downregulation in lipid biosynthesis in HCD. This suggests that lipid composition of the cell membrane may be one of the key factors contributing to the CDE. Lipids such as squalene synthase, HMGCS1 and cholesterol or fatty acids production enzymes were significantly enriched in EGE. Cell metabolic engineering upregulating such lipids or its precursors might render effective for reducing the CDE. Likewise, media supplementation with a synthetic cholesterol concentrates or lipid derivatives might be beneficial. This hypothesis is further explored in **Chapter two**.

The Unfolded Protein Response in rAAV production by Triple Transfection and Extended Gene Expression production

AAVs, in their quest to propagate, elicit a broad range of functions from their viral proteins, aiming to hijack the host cell machinery to facilitate their own replication, translation, packaging, and eventual egress from the cells. In response, the host cells strive to maintain their homeostasis counteracting these viral activities. A notable observation in this study is that apoptotic processes do not vary across the studied conditions, even though cell viability was clearly affected upon rAAV production[299]. Indeed, certain viral proteins required for rAAV production are well-known apoptosis inducers. AAV protein Rep78 activates caspase-3, leading to apoptosis[58]; and adenoviral proteins E4orf4 and E1A are responsible for p53-independent apoptosis[111,128]. The lack of alteration in processes related to apoptosis might be explained by the host cell response.

Cellular homeostasis was greatly disrupted upon rAAV production and differentially affected by the TT or EGE production methods. Particularly, the induction of endoplasmic reticulum (ER) stress over TT production is observed, which is characterized by an accumulation of misfolded or unfolded proteins that disrupt ER homeostasis[319]. Over evolution, cells have developed a complex network of signalling pathways aiming to detect and counteract ER stress, known as the Unfolded Protein Response (UPR). During rAAV production, primary stress-sensor pathways like PERK and ATF-6 were upregulated indicating an active UPR pathway. The UPR aim to restore normal ER function by halting protein translation, increasing the production of protein-folding chaperones, and promoting the degradation of misfolded proteins to improve cell survival. Indeed, in standard rAAV producing state, processes related to protein folding/refolding response in the ER were greatly enhanced in comparison to cell growth. Coherently, other studies on HEK293-based rAAV production reported that viral protein could trigger host cell stress responses, such as autophagy, ER stress, or apoptosis which were upregulated in the viral-producing state[299,320]. In contrast, the EGE method induced various metabolic modulations which may contribute to the recovery of cellular homeostasis. The EGE method showed an increase in E2 enzymes (ubiquitin-conjugating enzymes) which tag target proteins with ubiquitin for its degradation in the proteasome, indicating that degradation of misfolded proteins is favored. In addition, under the EGE condition, an increase in cytoplasmic chaperones, such as HSPA1, alongside a decrease in ER chaperones, such as calreticulin compared to TT production was observed. This suggests that the EGE method may lead to more efficient protein folding outside the ER, which are key functions of protein homeostasis system[321]. All these factors support the hypothesis that the EGE method seems to modulate the host cell metabolism to mitigate the ER stress, which is highly induced by rAAV production. These mechanisms contribute to counteract cell homeostasis disruption. Therefore, it could be considered that the EGE method exhibits a less pronounced UPR response compared to the TT method, which might offer advantages in terms of cell viability and productivity in rAAV production.

Mobilization of proteins to the cell nucleus: study of cellular importins

The study of importins is another critical factor in understanding the metabolic reprogramming of host cells in rAAV production methods. In the proteomic study, different importin groups were identified: mitochondrial importins, which role in energetic requirements has been previously discussed, and cellular importins, which transport proteins from the cytoplasm into the cell nucleus. In this work, a downregulation of cellular importins at HCD has been identified, similarly described by Lavado *et al*.[316]. However, novel findings in this study indicate that this decrease can be overcome by the EGE production method. This suggests that the EGE protocol might be more

efficient in mobilizing cellular proteins into the nucleus. In particular, the importin subunit α2 (KPNA2), which is a component of the nuclear pore complex (NPC) is highly upregulated in the EGE method of rAAV production (72 hpt) compared to the TT. To note, nuclear import of rAAV capsids has been reported to be mediated by host cell machinery using nuclear localization signals (NLS) of the N-terminal phospholipase A2 of VP1 capsid protein[322]. Specifically, the importin α/β heterodimer recognizes the NLS and facilitates nuclear import through the NPC[145,323]. Therefore, it is hypothesized that the EGE protocol might be favoring the canonical pathway of nuclear entry and rAAV capsid proteins among them. Consequently, suggesting that the enhanced rAAV production in the EGE method compared to TT is in part favored by more capsid availability in the cell nucleus. This observation may guide future host cell engineering to improve capsid availability in the cell nucleus which may be limiting the packaging step, currently known as one of the main limitations in rAAV production process.

Cell cycle modulation in rAAV production by Triple Transfection and the Extended Gene Expression methods

To date, the specific effects of PEI-mediated transfection on cell cycle regulatory mechanisms have not been thoroughly investigated. However, viral proteins essential for rAAV production can elicit diverse and even opposite responses in host cells, affecting cell cycle modulation. For example, the adenoviral protein E1A interacts with pRb, inhibiting it and favoring E2F activation, which promotes G1/S phase of the cell cycle progression[324,325]. On the other hand, Rep78 protein exerts antiproliferative effects on host cells, essentially causing a cell cycle arrest in the S phase through multiple pathways, including the activation of pRb[57]. Further, the adenovirus genes *E1A* and *E4orf6*, required for effective rAAV production, inhibit Cdc2 and promote degradation of cyclin A, keeping HEK293 cells in S phase[326]. These viral proteins are supplied either by co-transfection in plasmids or integrated in the HEK293 cell genome, and their coordinated expression within the host cell metabolism results in cell cycle alterations.

In the proteomic study, while no alterations were observed at 72 hpt in growth compared to the TT conditions, an increase was detected in the EGE compared to TT protocol. Subsequently, individual GO terms related to each phase of the cell cycle across the different rAAV production conditions were thoroughly studied.

Regarding the G1/S transition of the mitotic cell cycle, notable differences between the different rAAV production protocols were observed. During the rAAV production by TT, progression from G1 to S was hindered. In contrast, the EGE protocol showed a decrease in negative regulators of the

G1/S phase, thereby counteracting the downregulation of the S phase transition induced by rAAV production in the TT setting. Of note, it has been documented that AAVs modulate the host cell cycle to facilitate viral genome replication, occurring in the late S/G2 phase[327,328]. Prior studies have reported that the S-phase arrest is beneficial for both rAAV[328] and AdV production[329]. Based on the findings reported here, facilitating the transition into S phase could be similarly advantageous for rAAV production.

The mitotic cell cycle checkpoint is also impaired in the TT method, and this decrease is mitigated by the EGE protocol. Indeed, the negative regulation of the cell cycle is downregulated in EGE, suggesting that mitotic progression is favored. In this context, G2/M phase transition is increased in EGE. These observations in a context where ubiquitin-protein ligase activities involved in cell cycle are upregulated, might suggest that EGE protocol promotes faster degradation of these regulatory proteins, accelerating cell cycle progression. This is also supported by the increase in lipid biosynthesis in EGE which might be related to plasma membrane composition, indicating additional synthesis of phospholipids prior to cell division. In this context, the use of nocodazole, which arrests cells into the G2/M phase has been effective in enhancing rAAV production[330]. Interestingly, other works reported that transfection efficiency is strongly dependent on cell cycle phase at the time of transfection, with cells in the G2/M phase exhibiting greater transfection levels than cells at G1 phase[331]. A recent study showed that exogenous DNA is incorporated into the nucleus during telophase as the nuclear envelope reforms[332]. Therefore, nocodazole could also be used for enhancing transfection efficiency, improving the TGE method.

In conclusion, the detailed analysis of cell cycle dynamics during rAAV production reveals how the EGE method strategically modifies the cellular environment to support efficient viral particle production, compared to the TT method. These findings have the potential for targeted manipulation of the cell cycle.

Extracellular vesicle pathways modulation in rAAV production by Triple Transfection and the Extended Gene Expression methods

The EGE method in rAAV production is distinguished by its enhanced yield of rAAV particles in the cell supernatant fraction, outperforming the yields obtained through standard TT in batch that rely on cell lysis for harvesting. This study demonstrates that the EGE method significantly increases the fraction of rAAV secreted into the extracellular fraction, achieving 49% of rAAV egress with EGE compared to 35% of rAAV9 detected in supernatant with the standard TT approach, consistent with levels reported in literature[152].

The characteristic of rAAV egress was first characterized and exploited by Lock *et al.* and Vandenberghe *et al*. in 2010[151,152]. This highlighted the potential of harvesting rAAV from the suspension culture media of HEK293 cells, offering substantial advantages in reducing purification costs by minimizing contamination with cell-derived material[210]. However, the efficiency of rAAV release is low and strongly dependent on the rAAV serotype and the production protocol. In this context, the application of the EGE method for rAAV production has gained significant interest, due to its efficiency in enhancing the amount of rAAV harvested from the supernatant fraction[214]. Understanding the mechanisms through which rAAV exits the cell can have significant implications for optimizing production methods. This includes the development of new production protocols or modifications to host cell lines.

Previous studies have identified that a fraction of rAAV vectors isolated from the conditioned medium are associated with exosomes, termed exo-rAAV[153,154]. These exo-rAAVs exhibit enhanced transduction efficiency compared to conventional rAAV vectors[155]. Notably, the AAV membrane-associated accessory protein (MAAP) has been shown to be a viral egress factor by promoting rAAV exit from the cells by interacting with the surface of extracellular vesicles[89]. However, the mechanism(s) by which rAAV exits the cell remain to be determined.

In the conducted proteomic study, biological processes related to vesicle transport pathways are upregulated upon rAAV production by the standard TT and even higher with the EGE method. This suggests a potential link between the EGE method and enhanced vesicular egress. Therefore, it is hypothesized that the study of preferentially used proteins overexpressed in the EGE method compared to the TT, could provide insights into specific rAAV egress pathways favoured by the EGE protocol. The upregulation of the extracellular vesicle pathway together with the identification of increased rAAV egress in EGE, may indicate that the cellular machinery used for extracellular vesicle (EV) formation and release is preferentially used to favour rAAV egress. Hence, EV pathways are explored towards its exploitation for rAAV production.

It has been reported that nonenveloped Parvoviruses hijack host secretory transport pathways to promote their propagation. The initial stages involve the engulfment of viral particles by COPII-coated vesicles trafficking to the Golgi from the ER. Matured virus particles are then transported to the plasma membrane via the secretory pathway and actin filaments, and are finally released[333]. In the proteomic study, COPII vesicle coating do not exhibit any significant alterations under the TT or EGE conditions. This suggests that COPII-mediated ER to Golgi transport might not be the predominant mechanism driving rAAV secretion in the EGE method. In contrast, intra-Golgi

mediated transport instead is increased together with an increase in microtubule formation and binding.

Membrane fusion dynamics may also help explaining physiological events related to vesicle trafficking. In the EGE protocol, membrane fusion processes were upregulated. However, even more rAAV at cell supernatant were present in EGE, the endocytosis pathways were not enhanced. This suggests that the membrane fusion pathways belong to secretory pathway, fusioning the organelles or MVB to the plasma membrane to release rAAV particles or microvesicles.

Among vesicular transport pathway mediating the transfer of molecules between various cellular compartments, the biogenesis and secretion of exosomes represent a specialized subset[317]. Exosomes are small, lipid-bound extracellular vesicles that originate from the endosomal system. Specifically, when a portion of the limiting membrane of an endosome invaginates, it engulfs cytosolic material, leading to the formation of intraluminal vesicles within what is termed a multivesicular body (MVB). When MVBs fuse with the plasma membrane, these intraluminal vesicles are secreted as exosomes to the extracellular environment. The loading of cargo into MVB has been reported to follow multiple pathways. At least three different routes have been described: the ESCRT pathway, the tetraspanin pathway and a route involving lipids such as sphingomyelin[317]. The ESCRT machinery is key in the biogenesis of multivesicular bodies (MVBs). The ESCRT-I complex plays a pivotal role in the initial recognition and sorting of ubiquitinated proteins into the MVB pathway. The elevated levels of VPS37B in the proteomic study suggest an enhanced ESCRT-I mediated sorting mechanism, indicating a preference for this route for cargo sorting into MVB during the EGE protocol of rAAV production compared to TT. Other mechanisms of exosome biogenesis influencing ESCRT-independent loading of MVB are based on tetraspanins. ESCRT-independent sorting by tetraspanin CD63 has been reported by others[334]. Here, only CD63 was upregulated of all tetraspanins in the EGE method compared to TT. Coherently, the CD9 tetraspanin overexpression increased exo-rAAV production in other studies[153]. These observations may indicate the requirement of tetraspanins in the rAAV secretion pathway via exosomes. Another route for loading the MVB with cargo to be later released in the form of exosomes is via ESCRT-independent and tetraspanin-independent pathways like the one triggered by the metabolization of ceramide[335]. Ceramides are determinant in the vesicular transport pathway, as they induce negative membrane curvature that leads to the formation of ILV. The expression of sphingomyelinase or sphingomyelin phosphodiesterase enzymes breaking down sphingomyelin to ceramide has been reported to contribute to the pathways sorting cargo into the different routes within the MVB. Here, the neutral sphingomyelinase (nSMase) SMPD4 was upregulated in EGE

production compared to TT, suggesting a contribution of ceramide-dependent pathway in exosome release during rAAV co-production.

The second step in exosome generation is the fusion of the MVB to the cell membrane, releasing its cargo to the extracellular media. Therefore, proteins involved in exosome secretion regulation were studied such as the families of Rab GTPases and synaptosomal-associated proteins (SNAP), which have been reported to be heavily involved[317,336,337]. Rab GTPases, serve as master regulators of intracellular vesicle trafficking. Their dynamic modulation ensures the fine-tuning of vesicle budding, mobility along actin and tubulin networks, and targeted fusion, thus ensuring timely and precise delivery of vesicular cargo. In the conducted study, an increase in Rab GTPases in the EGE method compared to TT was observed at 96 hpt. In addition, among SNAP, specifically, SNAP47 was upregulated after the ME and RT step. These proteins are integral components of the SNARE (soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptor) complex. This complex is essential for vesicle fusion events within cells, a fundamental process for the delivery of vesicular cargo to intended destinations.

The comprehensive analysis of secretory pathways conducted herein highlights key proteins which might influence rAAV egress pathway. Based on these observations, investigating the targeted vesicular transport processes could provide crucial insights into the potential interplay between rAAV and EV secretory pathways. This is explored in **Chapter two.**

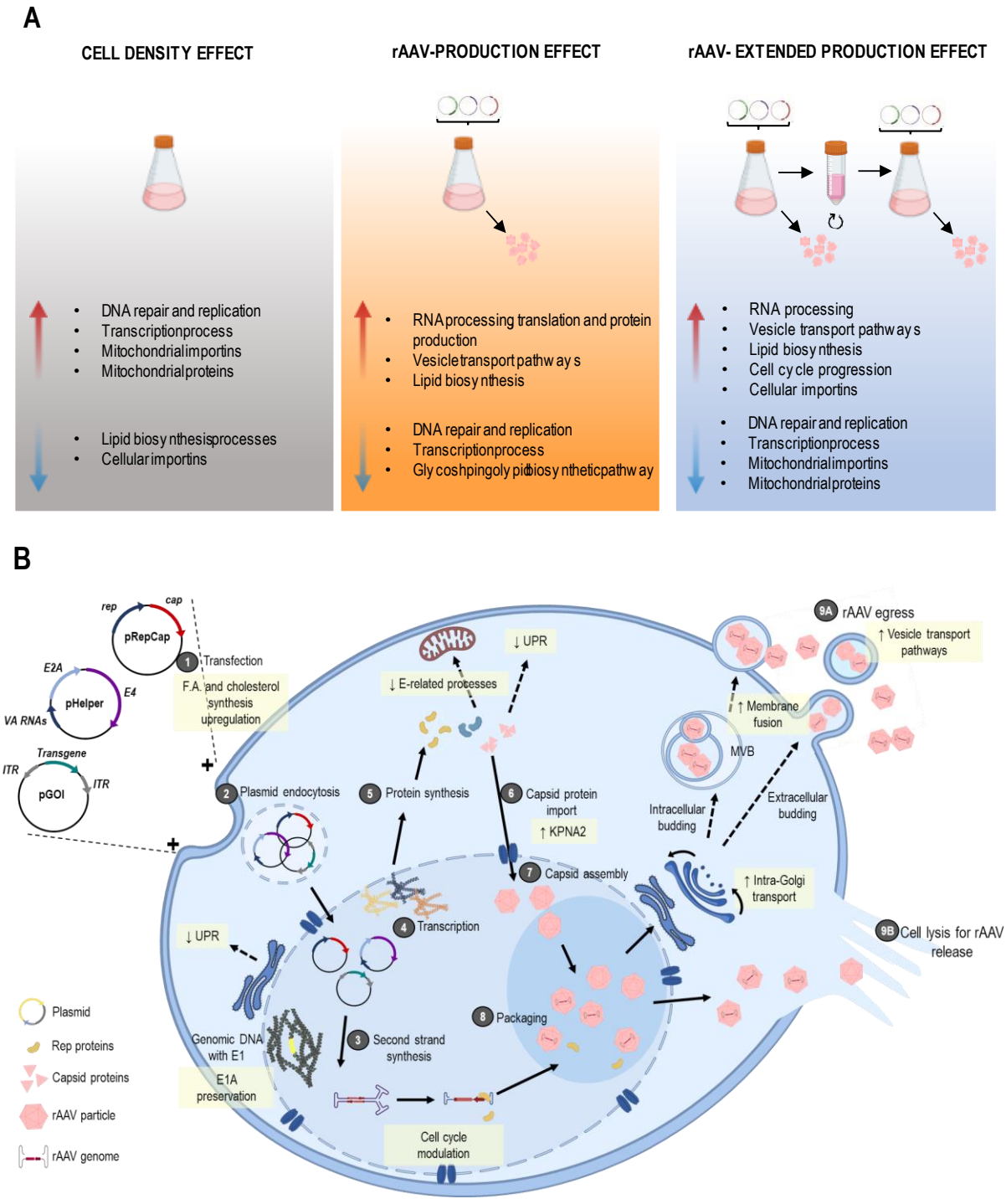


Figure 8. Summary of altered biological processes and proteins in the proteomic study. A. Summary of altered biological processes over time in growth condition (cell density effect). rAAV production effect alterations are related to effects present in rAAV TT production compared to cell growth. rAAV extended production effects are related to alterations present in EGE in comparison to TT rAAV production method. **B.** Altered biological processes and proteins in EGE compared to TT method for rAAV production in a cell biological context summarized in yellow boxes. Dashed lines indicate putative pathways hypothesized in this work. FA: fatty acids, MVB: Multivesicular bodies, UPR, Unfolded Protein Response.

CONCLUSIONS

The proteomic comparison on different rAAV production methodologies in HEK293 cells provides crucial insights into the cellular pathways affected during rAAV production. This study reveals a significant disruption in host cell homeostasis, predominantly characterized by an upregulation in the UPR and other stress responses, evidencing the importance of managing ER stress and protein homeostasis during vector production.

The EGE method compared to the standard TT production approach, demonstrates notable metabolic adjustments attributable to ME and RT steps. These differences include modulation of metabolic processes rendering a reduced mitochondrial activity, differential cell cycle modulation, upregulation of lipid biosynthetic processes and enhanced vesicle transport pathways, highlighting the potential of metabolic adjustments in host cell environment resulting in an enhanced rAAV production.

In conclusion, this comparative proteomic study identifies putative bottlenecks associated with standard TT method for rAAV production, serving as the groundwork for proposing strategic host cell modifications based on EGE-induced cellular reprogramming. These findings can guide new methods to enhance rAAV productivity through targeted modifications of the HEK293 cell line.

CHAPTER TWO

Metabolic engineering of HEK293 cells to enhance rAAV production and export

The contents of this Chapter generated the patent entitled: *Method for production of gene therapy vehicles in mammalian cells*, EP24159714.5.

INTRODUCTION

As discussed earlier, five gene therapy treatments based on recombinant adeno-associated virus (rAAVs) have been approved in the European Union for rare diseases, with the three most recent approvals occurring in 2022[20]. These approvals have contributed to the accelerated development of rAAV-based therapeutics, with over 200 clinical trials currently conducted worldwide and continuously increasing (ClinicalTrials.gov). Therefore, to meet future demand, the development of robust, large-scale rAAV manufacturing strategies is key to support the expanding field of rAAV-based gene therapies.

Currently, transient gene expression (TGE) in HEK293 cells by triple transfection (TT) is the most established method for rAAV production[184,199]. This approach enables rapid manufacture of viral vectors in sufficient quantities to support clinical development and commercial manufacturing, particularly for therapies targeting rare diseases. An important advantage lies in the ability to readily modify the specific rAAV product by changing plasmids coding for the gene of interest or by switching serotype through altering the capsid gene sequences. This flexibility facilitates early-stage process optimization and evaluation of rAAV therapeutic potential. However, the TGE approach for rAAV production faces several challenges that limit its large-scale application. First, transient transfection involves forming a complex between plasmid DNA and transfection reagents, resulting in a highly variable process. In addition, large quantities of high purity plasmids are required, significantly contributing to the overall cost of goods. Further, in TGE, unlike in Stable Gene Expression (SGE) approaches, expression occurs in a short time window, as plasmids remain episomal in the cells, limiting production timeframe.

Boosting transfection efficiency is key to achieve high protein expression to support rAAV production. However, the production by TT at large scale, where high cell densities (HCD) are desirable to increase viral vector productivity, is severely limited by the cell density effect (CDE)[190,281,287]. The CDE is a reduction in transfection efficiency and productivity in cells growing beyond the mid-exponential phase, and has been shown to negatively affect large-scale rAAV production[282–286]. Current rAAV production methods at large-scale using TT in suspension HEK293 cell lines are performed in WAVE bags or stirred tank bioreactors, reaching a maximal cell density for efficient transfection of 1 – 2 x 10^6 cells/mL and a maximal demonstrated capacity of 200 L[210,213,338]. In this setting, rAAVs are produced in batch mode, and typically harvested from the cell lysate fraction for full rAAV recovery, yielding up to 1 x 10^5 vg/cell[208,277].

Process intensification strategies, such as perfusion-based approaches have been implemented to increase rAAV manufacturing by increasing cell densities and viral vector productivities[210,286,339]. During perfusion cultures, cells are maintained in the bioreactor by filtration devices that enable simultaneous cell retention and medium exchange, which can be performed continuously or at discrete time points, preventing nutrient depletion while removing growth-inhibiting compounds. These methods offer a dual-approach strategy to improve rAAV manufacturing. First, media replacement strategies have been described to modulate the cell physiological condition associated with the CDE, proving valuable for achieving transient transfection at HCD for TT rAAV production[281,286]. Moreover, as reported by Grieger *et al*. in 2016, the perfusion operation mode enables a continuous harvest of the rAAV from the cell culture media, with minimal loss of the rAAV in the cell pellet fraction[89,151,152,210]. Harvesting only the supernatant fraction, exploiting the rAAV property of cell release avoids the introduction of additional contaminants in comparison to rAAV harvest via cell lysis as in batch production, thus simplifying downstream purification processes by reducing the presence of cell-derived contaminants[151,276,340–342]. However, the effectiveness of this strategy is limited by low rAAV yields at the supernatant fraction, dependent on the rAAV egress property, which is serotype-dependent[152,343]. To address these challenges, the application of the Extended Gene Expression (EGE) production strategy, previously described for viral particles has been identified as a promising approach for rAAV production enhancement[214,294]. The EGE protocol involves consecutive medium exchanges (ME), followed by re-transfection rounds (RT), extending the production timeframe and introducing more functional DNA. This approach has led to a remarkable three-fold increase in total rAAV yield harvested from the supernatant compared to TT production in batch and cell lysate recovery[214].

In **Chapter one**, the superiority of EGE method for rAAV production compared to TT has been studied, revealing that the EGE not only enhances the rAAV9 yield at harvest but also improves viral egress. Using a multiplexed quantitative proteomic analysis, the molecular dynamics of host cell response modulating the rAAV production pathway differently in EGE or TT were studied. This analysis pinpointed several key molecular targets for optimization through metabolic engineering techniques. Understanding the molecular dynamics in the EGE context is essential to explore rational modifications to the host cell line, aimed at increasing rAAV recovery in the cell supernatant.

In this work, the mitigation of the CDE effect, negatively impacting TT rAAV production in HEK293 cells was first targeted aiming to maintain transfection efficiency at HCD. The proteomic analysis presented in **Chapter one,** revealed that cells at HCD experienced a downregulation in lipid biosynthetic pathways[316,344]. However, following the EGE protocol, enzymes such as squalene

synthase, Hydroxymethylglutaryl-CoA Synthase 1 (HMGCS1) or fatty acids production enzymes were significantly enriched. Given that ME has been extensively reported to mediate efficient transfection at HCD[281], in this chapter, lipid supplementation strategies were tested as potential modulators of the CDE, evaluating their performance for transfection efficiency at HCD for rAAV production.

Furthermore, efforts were focused on enhancing the recovery of rAAV from the extracellular fraction. Previous studies have documented that a portion of rAAV vectors are naturally released associated with extracellular vesicles (EV) into the supernatant of the cell culture media[345]. Indeed, cells release a range of membrane-bound vesicles, with sizes varying from 20 nm to 1 μm in diameter, commonly identified as exosomes or microvesicles, here collectively referred to as extracellular vesicles or EVs. These vesicles are capable of encapsulating diverse macromolecules, including proteins, nucleic acids, and viruses. Notably, rAAVs associated with exosomes are specifically termed exosome-associated rAAV (exo-rAAV)[151,152,345]. Research has demonstrated that exo-rAAVs outperform conventional rAAV vectors in several aspects: they have greater transduction efficiency, exhibit resistance to neutralizing antibodies[346–348], and show higher efficiency transducing hard-reaching tissues efficiently crossing the blood-brain barrier[155] or the retina inner membrane[349]. Recent studies have linked the mechanism of rAAV egress from cells to their association with EVs, specifically through the AAV membrane-associated accessory protein (MAAP). This protein has been identified as a factor promoting rAAV egress from the cells by interacting with the surface of EVs[89].

In the proteomic study in **Chapter one**, it was observed that vesicle transport pathways were upregulated in the EGE method compared to the TT rAAV production method. This change in extracellular vesicle pathway dynamics, along with the observation that EGE is associated with higher rAAV export to the cell supernatant, suggests a potential link between the EGE method and enhanced vesicular egress during rAAV production which is explored in this work. Further, the study of the preferential use of certain proteins involved in the exosome generation pathway in the EGE method compared to TT, provides insights into the favoured rAAV egress pathway within the EGE protocol. These findings highlight specific mechanisms that may enhance the efficiency of rAAV export through vesicular transport, as represented in **Figure 1**.

In **Chapter one**, vesicle transport proteins upregulated in the EGE protocol compared to TT, were related to exosome biogenesis and secretion pathway. Specifically, exosomes are small, lipid-bound extracellular vesicles that originate in the endosomal system and represent a specialized subset in the vesicular transport pathway[317]. They origin when a portion of the limiting membrane

of an endosome invaginates, engulfing cytosolic material, leading to the formation of intraluminal vesicles within what is termed a multivesicular body (MVB). When MVBs fuse with the plasma membrane, these intraluminal vesicles (ILVs) are secreted as exosomes into the extracellular environment. The loading of cargo into MVBs follows multiple pathways[317]. At least three routes have been described: the ESCRT pathway, the tetraspanin pathway and a lipid-based route involving lipids such as sphingomyelin[317]. First, the increased expression of VPS37B in the EGE condition suggests the involvement of an ESCRT-I mediated sorting mechanism. In addition, CD63 following a ESCRT-independent sorting by the tetraspanin pathway was upregulated in the EGE protocol[334]. Regarding the lipid route, enzymes like sphingomyelinase or sphingomyelin phosphodiesterase, which break down sphingomyelin into ceramide, contribute to the sorting of cargo into different routes within the MVBs[335]. Ceramides, inducing negative membrane curvature, lead to the formation of ILVs. The neutral sphingomyelinase SMPD4, showed increased expression in the EGE production setting. In addition, proteins involved in the fusion of the MVB to the cell membrane, releasing cargo to the extracellular media, such as the families of Rab GTPases and synaptosomal-associated proteins (SNAP), heavily involved in this process, were studied[317,336,337]. Rab32, was identified with the highest number of proteotypic peptides among more than 46 different quantified Rab proteins, and SNAP47, showing the highest upregulation in EGE compared to TT and non-transfected conditions. Subsequently, rationally targeted proteins, namely SNAP47, Rab32, VPS37B, SMPD3 and CD63 were selected as potential candidates for further exploration of its role in rAAV production pathway (**Figure 1**).

Overall, this work studies effective modulators of the TGE method by TT for rAAV production at different levels. First, lipid supplementation strategies are developed to study the CDE modulation, desirable for large scale application of the TT method. Further, rational modifications through metabolic engineering of the EV pathway are studied, aimed at increasing the rAAV harvest at the extracellular fraction. Such an approach, could inform future host cell engineering strategies for rAAV production, enabling to increase extracellular rAAV production yields, thereby facilitating the downstream process.

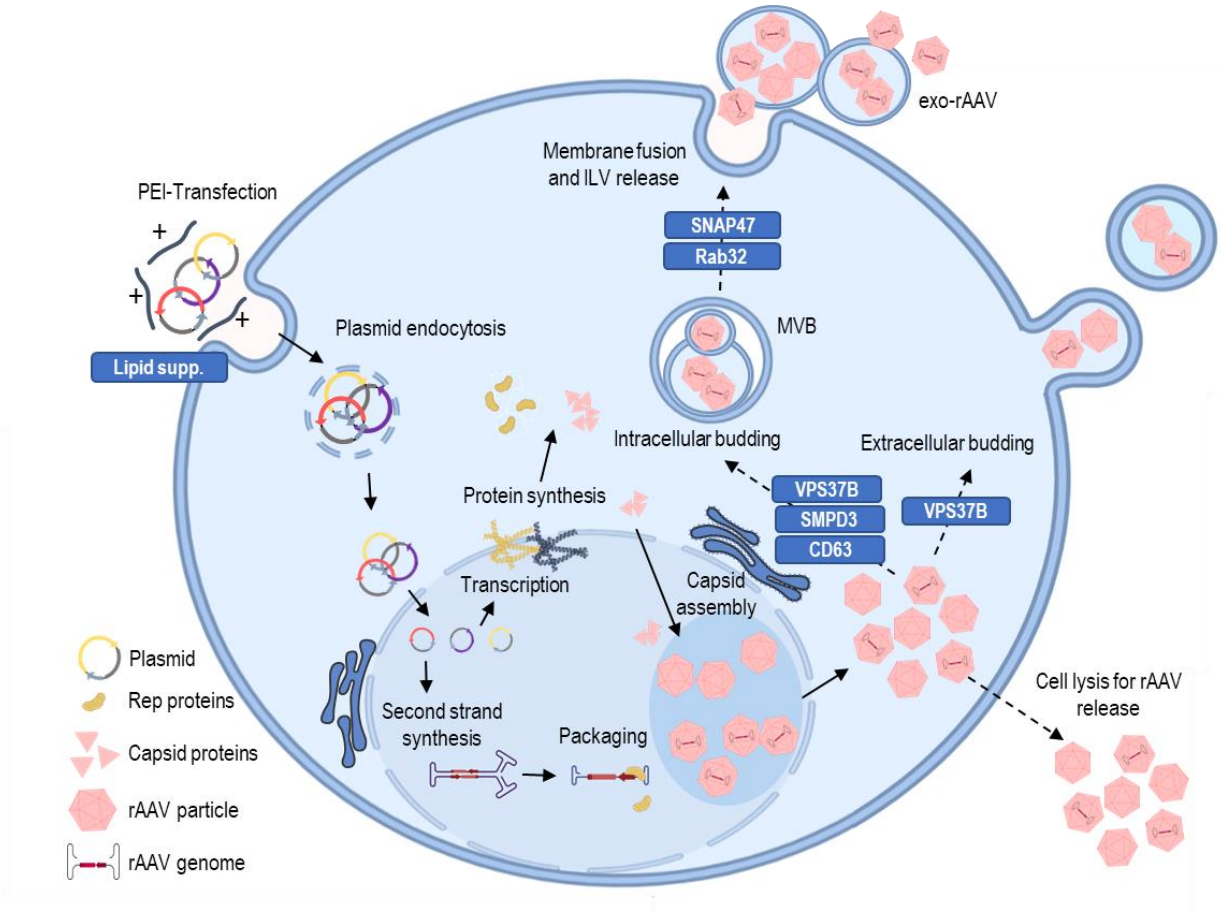


Figure 1. Schematic representation of the rAAV production pathway by triple transfection, with hypothesized EV-targeting pathways depicted. The schematic illustrates the steps involved in the rAAV production process, including transient transfection, replication, assembly, and packaging within host cells. Typically, in the standard triple transfection method, rAAV particles are harvested through cell lysis. However, rAAV vectors can also be collected from the supernatant, where a portion is naturally released in association with exosomes (exo-rAAV). This figure depicts the proposed targets of the EV pathway, which may modulate the egress of rAAV, indicated by dashed lines: proteins SNAP47 and Rab32 involved in exosome fusion and release, SMPD3, CD63 and VPS37B which participate in intracellular budding process. In addition, lipid supplementation strategies are proposed for CDE modulation. CDE, cell density effect; ILV, intraluminal vesicles; MVB, multivesicular bodies; PEI, polyethyleneimine.

MATERIALS AND METHODS

Mammalian cell line and culture conditions

The cell line used in this work is a serum-free suspension-adapted HEK293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from the National Research Council of Canada (NRC, Montreal, Canada). The CHO cells used is a serum-free suspension-adapted, CHO-S cell line (ExpiCHO-S™ ThermoFisher Scientific, San Jose, CA, USA).

Cell culture media for HEK293 cells were HyCell™ TransFx-H (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA) and 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, ThermoFisher Scientific, San Jose, CA, USA). FreeStyle™ F17 Expression media (Gibco, ThermoFisher Scientific, San Jose, CA, USA) supplemented with 8 mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA), 0.1% Pluronic™ F-68 Non-ionic surfactant (Gibco, ThermoFisher Scientific, San Jose, CA, USA) and IGF-1 (91590C, Sigma-Aldrich, St. Louis, MO, USA) at a final concentration of 50 µg/L. For CHO-S cells, cell culture media was FreeStyleCHO™ optimized in-house with supplementation of 8 mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA) and 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, ThermoFisher Scientific, San Jose, CA. USA).

Suspension cell cultures were maintained in exponential growth phase in disposable polycarbonate 125 mL flasks with vent cap (Corning, NY, USA) at 37ºC, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Cell concentration and viability were determined using the NucleoCounter® NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

When required, according to the planned experiments, culture media was supplemented with lipids: Ceramide at 2 mg/L (860052 P, Avanti, Sigma-Aldrich, St. Louis, MO, USA) was dissolved in dimethyl sulfoxide 99% (DMSO; W387520, Sigma-Aldrich, St. Louis, MO, USA); cholesterol (SyntheChol®, Sigma-Aldrich, St. Louis, MO, USA) and fatty acids (F7050, Sigma-Aldrich, St. Louis, MO, USA) were both diluted in PBS (HyClone™, GE Healthcare, Chicago, IL, USA). These supplements were added to the culture media at concentrations specified for each experimental condition.

Plasmid design and generation

Plasmids used for standard triple transfection were kindly provided by Dr. Fàtima Bosch from Centre for Animal Biotechnology and Gene Therapy (CBATEG), at Universitat Autònoma de Barcelona (UAB, Bellaterra, Spain). These included plasmids pHelper, pGG2_CMV_GFP (referred to as pGFP_ITR hereafter), and pRep2Cap9. The pHelper plasmids codifies for the adenovirus elements necessary for rAAV replication and packaging[214]. The pGFP_ITR plasmid encodes for the enhanced GFP gene under the control of the cytomegalovirus (CMV) promoter and is flanked by AAV2 inverted terminal repeats (ITRs) to enable its packaging into rAAV capsids. The plasmid pRep2Cap9 encodes the Rep proteins from AAV2 and capsid proteins from AAV9.

For transfection control, pMock plasmid was used. This plasmid, devoid of any mammalian promoter or coding DNA sequence (CDS) was engineered by ligating the pGag::eGFP backbone[350,351].

To monitor transfection efficiency when required, pGFP plasmid, which encodes for the Green Fluorescent Protein (GFP), was used as a reporter gene. It was constructed by PCR amplification of pGagGFP backbone using primers containing EcoRI restriction digestion site as shown in **Table 1** to cut Gag polyprotein sequence. PCR product was digested with EcoRI-HF® (R3101S, New England Biolabs (NEB), Ipswich, MA, USA) and remaining product was re-ligated using T4 DNA ligase (NEB, Ipswich, MA, USA).

Plasmids pSNAP47 and pRab32 were generated in house. Briefly, *snap47* and *rab32* genes (Nucleotide NCBI Reference Sequence: NM_053052.4 and NM_006834.5 respectively) were obtained as a gBlock at IDT (Integrated DNA Technologies, Leuven, Belgium). They were amplified by polymerase chain reaction (PCR), and products were checked by gel electrophoresis and then digested with EcoRI-HF® (R3101S, NEB, Ipswich, MA, USA) and NotI-HF® (R3189S, NEB, Ipswich, MA, USA) for pSNAP47 or EcoRI-HF® and KpnI (ER0521, ThermoFisher Scientific, San Jose, CA, USA) for pRab32. Fragments were cloned into pGFP vector previously digested with the same mentioned restriction enzymes using T4 DNA ligase (NEB, Ipswich, MA, USA).

The shRNAs were designed using siRNA design tool by IDT. Here, they were obtained as ssDNA Ultramer™ (IDT, Leuven, Belgium), flanked by regions with sequence complementarity to a backbone containing U6 promoter which was PCR amplified[344]. The PCR fragment with the ssDNA were incubated with Hi-Fi Master Mix 2x (NEB, Ipswich, MA, USA) for 1 hour at 50ºC using a molar ratio of 1:5 or 1:10. The siRNA sequences are described in **Table 1.**

Plasmids pCD63 and pSMPD3 were obtained from Genescript. Shortly, transcript variant NM_018667.4 for *smpd3* gene and NM_001257389.1 for *cd63* gene were cloned into pcDNA3.1 +/C-(K)DYK plasmid. Plasmid pVPS37B was constructed by Genescript using cDNA (NM_024667.3) and cloned into pcDNA3.1 +/C-(K)DYK plasmid.

Plasmids were amplified in a chemicompetent *E. coli* DH5α strain produced in house growing in LB medium (Conda, Madrid, Spain) supplemented with 10 µg/ml kanamycin (Sigma-Aldrich, St. Louis, MO, USA) or 100 µg/ml ampicillin (Sigma-Aldrich, St. Louis, MO, USA) depending on the *E. coli* antibiotic resistance present on each plasmid. Plasmid purification was carried out using the Endofree Plasmid MDI kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Quantity of isolated DNA was assessed using Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, San Jose, CA, USA). Plasmids were confirmed by Sanger sequencing (Servei de Genòmica, UAB, Bellaterra, Spain).

Table 1. List of primers and sequences used. Primers were designed avoiding possibilities of self-dimerization and primer-dimer formation. siRNA sequences were designed using IDT siRNA selection tools and ThermoFisher design tools.

Name	Objective	Sequence (5'-3')
Fwd_GFP_EcoRI	Clonning of pGFP plasmid vector from pGagGFP	CTACAGGAATTCGATCCACCGGTCGCCACC AT
Rev_GFP_EcoRI		GGAGGCTGGATCGGTCCGGT
shCD63	siRNA sequence for shRNA plasmid construction	GCAAGGAGAACTATTGTCTTA
shCNT		CCTAAGGTTAAGTCGCCCTCG
shVPS37B.1		GATTCCTTCATTGATGTCTATCAGA
shVPS37B.2		ATAAAGAAGACCAAATTAG
Fwd_GFP_qPCR	qPCR for rAAV titer determination	AGCTGGACGGCGACGTAA
Rev_GFP_qPCR		GTGCAGATGAACTTCAGGGTCA
GFP gBlock	gBlock used as standard in qPCR	AGCTGTTCACCGGGGTGGTGCCCATCCTG GTCGAGCTGGACGGCGACGTAAACGGCCA CAAGTTCAGCGTGTCCGGCGAGGGCGAGG GCGATGCCACCTACGGCAAGCTGACCCTG AAGTTCATCTGCACCACCGGCAAGCTGCCC GTGCCCTGGCCCACCCTCGTGAC

Transient transfection protocol

Transient transfection of HEK293 suspension cells was performed at a cell density of 2 x 10^6 cells/mL. For HCD experiments, transfection was performed at 6 x 10^6 cells/mL. As transfection reagent, 25-kDa linear polyethilenimine (PEIpro®, Polyplus SA, Illkirch-Graffenstaden, France) was used according to planned experiments. Transfections were carried out using a final DNA concentration of 1 µg/mL with a DNA:PEI ratio of 1:2 (w/w).

Transient transfection of CHO-S cells was performed at a cell density of 1.2 x 10^6 cells/mL using FectoPro® as transfection reagent (Polyplus SA, Illkirch-Graffenstaden, France) at a final concentration of 0.75 µg/mL of DNA with a DNA:FectoPro® optimized ratio of 1:2 (w/w) . After transfection 0.5 µl/mL of FectoPro® Booster was added directly to the cell culture.

For transient transfection, DNA:PEI complexes were formed by adding PEI to plasmid DNA diluted in supplemented HyCell or FreeStyleCHO media for HEK293 or CHO-S cells respectively. Transfection mixture of 10% of the total cell culture volume to be transfected was prepared and vortexed for 10 seconds. Then, PEI was added and vortexed three times for 3 seconds, incubated for 15 minutes at room temperature and then added to the cell culture. According to experimental requirements, transient transfections were performed with 2 mL total culture volume in six-well plate suspension format or 20 mL total culture volume in shake flasks.

rAAV production

For rAAV production, transient transfection (TT) was performed in HEK293 cells using an equal plasmid molarity of pHelper, pGFP_ITR and pRep2Cap9. Additional tested plasmids were added to the TT plasmids using equimolar concentrations. Total DNA was normalized to 1 µg/ml using pMock when necessary.

For rAAV production by TT method, cell culture was harvested 72 hpt or 96 hpt. For EGE method, media exchange (ME) was performed 48 hpt by centrifuging the cell culture at 300 x g for 5 minutes. The cell supernatant was collected, and the cell pellet was gently resuspended with fresh media. After that, re-transfection (RT) was performed according to previous transient transfection methodology. Cell culture was harvested 96 hpt. All conditions were performed in triplicate.

For rAAV characterization, cell culture samples were taken at 48, 72, and 96 hpt according to rAAV production method and experimental requirements. For cell supernatant collection, cell culture was centrifuged at 1000 x g for 10 minutes. For cell lysate collection, whole cell culture was treated with 1x cell lysis solution (500 mM Tris, 1% Triton X-100 (v/v), 20 mM $MgCl_2$ at pH 7.5) and incubated

at 37ºC, 130 rpm, for 1 hour. Lysate mixture was centrifuged at 4000 x g for 20 minutes and supernatant was collected as cell lysate fraction. All fractions were stored at -80ºC until analysis was performed.

Flow cytometry

The percentage of GFP positive cells was assessed using Cytoflex (Beckman Coulter Inc., Brea, CA, USA) at Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria (UAB, Bellaterra, Spain). First, SSC-H vs FSC-A and FSC-A vs FSC-H density plots were generated to gate individual cell population. Then, GFP positive population was assessed in Y585-PE-A vs GFP B525-FITC-A density plot. Gating was manually adjusted with HEK293 non-transfected cells as a control. Data was processed using CytExpert v.2.3 software (Beckman Coulter Inc., Brea, CA, USA).

Harvest of extracellular vesicles

The harvesting protocol of extracellular vesicles (EV) was adapted from Schiller *et al.*, 2018[153]. Briefly, cell media of 20 mL shake flask culture was harvested and centrifuged at 300 x g for 5 min and at 1000 x g for 10 min sequentially. The supernatant fraction was then centrifuged at 17,700 g for 1 h at 4ºC. Then, the supernatant was centrifuged again at 100,000 x g for 1.5 h at 4ºC in Beckman Optima L100P centrifuge (Beckman Coulter Inc., Brea, CA, USA) and the pellet was resuspended in 200 µL of supplemented F17 cell culture media. Samples were kept at -80 ºC. Centrifugations were performed for each experimental condition in triplicate.

Quantitative polymerase chain reaction (qPCR) for rAAV quantification

For the determination of physical vector genome titer of rAAV (vg), 5 µL of the cell lysate fraction or the cell supernatant fraction from each sample was added to DNase I (0.066 mg/mL, 10104159001, Roche, Basel, Switzerland) prepared in DNase I buffer (0.16 mM $CaCl_2$, 6.6 mM $MgCl_2$, 17.3 mM Tris-HCl, 0.066 % Pluronic F68, pH 8.0). This mixture was incubated at 37 ºC for 16 hours. Then, 4 µL of 30 mM EDTA (AM9260G, Invitrogen, Carlsbad, CA, USA) in 10 mM Tris-HCl, pH 8.0) was added to each tube and incubated at 70ºC for 30 minutes to inactivate DNase I. After, 5 µL of Proteinase K (EO0491, ThermoFisher Scientific, San Jose, CA, USA) was added, incubated at 55ºC for 2 hours and inactivated at 95ºC for 15 minutes. The quantification of viral genomes was conducted using quantitative PCR (qPCR). This reaction was set up using iTaq Universal SYBR Green Supermix (1725124, BioRad, Hercules, CA, USA) and a set of primers specific to a region within the GFP gene (**Table 1**). Thermocycling profile was based on

manufacturer's recommendations in a CFX384 Touch Real-Time PCR Detection System (**Table 2**). For the establishment of a standard curve, a gBlock (IDT, Leuven, Belgium) containing the targeted GFP region of a known concentration was used (sequence shown in **Table 1**). A serial dilution of this standard, ranging from 10^8 to 10^2 copies/µL, was prepared and amplified by qPCR. The concentration of rAAV samples was the determined by interpolating their amplification data onto the standard curve. Additionally, an internal rAAV1-CMV-GFP reference was amplified alongside the samples in all the quantifications.

Table 2. Conditions of cycling temperatures and time for qPCR and RT-qPCR.

Reaction	Cycles (temperature and time)
qPCR for rAAV quantification and RT-qPCR	95ºC, 3 min + 40 x [95ºC, 10 sec + 60°C, 30 sec] + 95ºC, 10 sec + 65 to 95°C, 0.5ºC/5 sec + 4°C hold

To quantify the accumulated rAAV titer in the supernatant (SN) (extracellular media), **Equation (1)** was used. To quantify the rAAV accumulated titer in the whole cell culture, the cell lysate (Lys) fraction, **Equation (2)** was used. The total accumulated rAAV in the supernatant (Acc. rAAV in SN) was determined by adding the titer of supernatant fraction at harvest (96 hpt) to the titer of the harvested supernatant fraction at 48 hpt. Similarly, the total accumulated rAAV in the cell lysate fraction (Acc. rAAV in Lys) was determined by adding the titer of cell lysate fraction at harvest (96 hpt) to the titer of the harvested supernatant fraction at 48 hpt.

$$\text{Acc. rAAV at SN (vg)} = \text{rAAV titer in SN 96 hpt} \times V + \text{rAAV titer in SN 48 hpt} \times V \quad \textbf{(1)}$$

$$\text{Acc. rAAV at Lys (vg)} = \text{rAAV titer in Lys 96 hpt} \times V + \text{rAAV titer in SN 48 hpt} \times V \quad \textbf{(2)}$$

To quantify the specific rAAV production in each cellular fraction, normalized rAAV titer at harvest in either SN or Lys fraction by the viable cell density (VCD) at the time of harvest was calculated, using **Equation 3** as follows:

$$\text{rAAV titer in SN or Lys (vg/cell)} = \frac{\text{rAAV titer in SN or Lys (vg/mL)}}{\text{VCD at harvest (viable cells/mL)}} \quad \textbf{(3)}$$

To determine the rAAV egress rate (referred as Esn), the proportion of rAAV present in the supernatant fraction relative to the total rAAV was calculated using **Equation (4)**:

$$\text{Esn (\% SN)} = \frac{\text{rAAV titer in SN}}{\text{rAAV titer in Lys}} \times 100 \quad \textbf{(4)}$$

Extracellular vesicle quantification and characterization

Nanoparticle tracking analysis (NTA) was performed with a NanoSight® LM20 device (NanoSight Ltd., Amesbury, UK) at the Institut de Ciència de Materials de Barcelona (ICMAB, CSIC, Bellaterra, Spain), equipped with a neutral density filter for total particle by light scattering. Briefly, samples were injected being previously diluted 1:20 and three independent analyses were carried out. Three video recordings of 60sec length were made for each sample. Subsequently, particles were identified and tracked by their Brownian motion at room temperature. Data were analyzed with NanoSight® NTA 3.4 software.

Dynamic light scattering (DLS) experiments were performed using a Zetasizer Nano ZS instrument (Malvern instruments, Malvern, UK) with a He/Ne 633 nm laser at 173°. The hydrodynamic diameter, particle size distribution in volume, derived count rate (dCR) and polydispersity index (PDI) were calculated with cumulative fit correlation at 25°C and 0.887 cP or 2.448 cP. Briefly, 50 µL of sample was placed in disposable plastic cuvettes (UV-Cuvette micro, BRAND GMBH, Germany) followed by automated experimental data collection. Technical triplicates with 12 scans of 10 s were performed in each independent measurement.

rAAV9 in vitro potency assay

HEK293 cells were seeded at a concentration of 1 x 10^5 cells/well in a 96-well plate and incubated for 1 hour at 37ºC until cells were adhered. Then, rAAV9 samples were normalized to 1.2 x 10^{10} vg/mL and added to each well at different dilutions ranging from MOI 500 to 4000 in triplicate. wt AdV2 was kindly provided by Theriva™ Biologics (Parets del Vallès, Spain) and used at a multiplicity of infection (MOI) of 30 to enhance rAAV transduction efficiency[352]. Cells were harvested 24 hours post-transduction (hptd) and the percentage of GFP-positive cells was assessed using Cytoflex (Beckman Coulter Inc., Brea, CA, USA) at Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria (UAB, Bellaterra, Spain). First, SSC-H vs FSC-A and FSC-A vs FSC-H density plots were generated to gate individual cell population. Then, GFP-positive population was assessed in Y585-PE-A vs GFP B525-FITC-A density plot. Gating was manually adjusted with HEK293 non-transduced cells as a control. Samples to be analyzed were concentrated 2-fold before analysis. 5000 events were analyzed per sample at a flow rate of 70 µL/min. Results were analyzed with the CytExpert v.2.3 software. The TU was calculated using the following expression (**Equation 5**):

$$\text{Transducing units (TU)/mL} = \frac{\text{GFP} \times \text{N} \times \text{FD} \times 1000}{\text{V}} \quad \textbf{(5)}$$

where GFP is the number of GFP-positive cells per well, N is the number of cells per well at the time of transduction (corresponding to $1x10^5$ cells per well), FD is the fold dilution of vector sample used for transduction at different MOIs, V is the volume of diluted vector sample added into each well for transduction in microliters[352,353].

Dot Blot analysis

For Dot Blot, the extracellular fraction from each sample was loaded (50 µL) on nitrocellulose membrane, air dried and incubated for 1 hour in blocking buffer (1x PBS, 0.1% (v/v) Tween-20 with 5% (w/v) non-fat milk) followed by 1 hour incubation at room temperature with the corresponding primary antibody at room temperature with gentle shaking. Primary antibody used was rabbit anti-CD63 antibody (ab134045, Abcam UK, 1:1000). Membrane was washed three times for 5 minutes on orbital shaker with wash buffer (1x PBS, 0.1% (v/v) Tween-20). After primary incubation, a secondary incubation was performed using HRP-linked anti-rabbit (A3562, Sigma-Aldrich, St. Louis, MO, USA) antibody at 1:2500 in 2.5% (w/v) non-fat dry milk 1x PBS 0.1% (v/v) Tween-20 for 1 hour at room temperature. Proteins were visualized using Clarity Western ECL Substrate solution (BioRad, Hercules, CA, USA) using ChemiDoc MP Imaging System (BioRad, Hercules, CA, USA). For Dot Blot analysis, image was inverted, background signal was subtracted and integrated cell density was quantified. A standard curve using purified extracellular vesicles of known concentration was run in parallel. The concentration of extracellular vesicles in the samples were determined by interpolating its integrated density data in the standard curve.

Optimal combination of candidate genes by Design of Experiments (DoE)

To optimize the viral egress rate (Esn) and specific production (Psp), a Design of Experiments (DoE) strategy specifically employing a Simplex Design was used to determine the optimal ratios of various plasmids (pSMPD3, pSNAP47, pVPS37B) impacting these responses. The DoE responses were quantified using **Equation 3** Tto calculate the specific production in the cell lysate (Psp; vg/cell in Lysate) and **Equation 4** for viral egress rate (Esn; percentage of rAAV9 in the SN).

The impact of the different plasmids on each response was assessed by varying their proportions in the transfection mixture, ensuring that the combined total for each experimental run equaled 1, which was set at 9×10^{10} copies/mL. The specific proportions and combinations for each plasmid are detailed in **Table 3**. For each experimental run, the proportions of each plasmid were calculated relative to their 100% concentrations (in µg/µL) and different combinations of the three genes were prepared. Overall, the simplex-related mixture accounted for 9×10^{10} copies/mL of plasmid DNA, and the three additional plasmids for rAAV production (pHelper, pGFP_ITR and pRep2Cap9)

represented 1.8 x 10^{10} copies/mL to the total DNA concentration. To achieve a total DNA concentration of 1 µg/ml, pMock was used.

Simplex experimental results were fitted to a polynomial model described as follows (**Equation 6**):

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i \cdot X_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} \cdot X_i \cdot X_j \qquad (\mathbf{6})$$

where Y is the response variable under consideration (Esn (% of viral egress rate) or Psp (vg/cell at lysate) at harvest), β_0 is the model intercept term, β_i is the linear coefficient, β_{ii} is the interaction coefficient, and the coded factors X_i and X_j are the independent variables of study, *k* is the number of independent variables.

Statistical analysis and graphs facilitating model interpretation were performed using the Design-Expert® Software (Stat-Ease, Inc). The model coefficients were determined using Type III partial sum of squares. This allows for isolation of each variable's effect in the presence of other variables (**Table 4**). The fit of the model was quantified using the coefficient of determination (R^2), and the lack-of-fit (LOF) test was employed to evaluate the difference between experimental results and the pure error of the fitted equations. The overall significance of the model was determined using ANOVA F-tests, with each coefficient's significance assessed via the corresponding t-test. *p*-values of 0.05 and 0.1 were considered as statistically significant, corresponding to confidence levels of 95% and 90%, respectively.

Cryo-Transmission electron microscopy (Cryo-TEM)

A 2-3 µL amount of sample was blotted onto holey carbon grids (Quantifoil Micro Tools, Großloebichau, Germany and Micro to Nano, Haarlem, Netherlands). The samples were subsequently plunged into liquid ethane at −180°C using a Leica EM GP cryo workstation and observed in a JEM-2011 electron microscope (JEOL Ltd., Tokyo, Japan) operating at 200 kV. During imaging, samples were maintained at −173°C, and pictures were taken using a CCD-multiscan camera (Gatan Inc., Pleasanton, CA, USA).

Statistical analysis

All data were represented as mean ± standard deviation. Unless stated, one-way analysis (ANOVA) was used for all statistical analysis in this study. GraphPad was used to conduct statistical analysis. *p* values less than 0.05 were considered statistically significant.

RESULTS

Lipid supplementation strategy to improve transfection efficiency at high cell density

To mitigate the Cell Density Effect (CDE) that reduces transfection efficiency at High Cell Densities (HCD) and thus limits the efficiency of the triple transfection (TT) method for rAAV production, a lipid supplementation strategy is proposed. This approach is based on proteomic analysis findings detailed in **Chapter one**, that indicate a deficiency in fatty acid synthase and cholesterol-related synthetic pathways at HCD. By supplementing specific lipids, this strategy aims to enhance these pathways, potentially improving transfection efficiency and overall rAAV yield.

To this end, varying concentrations of fatty acid and synthetic cholesterol supplements were tested alone or in combination. Specifically, fatty acids were added at 0.5 pg/cell or 0.25 pg/cell for high and low concentrations, respectively. Similarly, cholesterol was added at 10 pg/cell and 5 pg/cell for high and low concentrations, respectively. These supplements were added to the cell culture 4 hours before transfection (hbt).In addition, a positive control condition involving a ME 4 hbt was included, which has been previously reported to mitigate the CDE[281]. The cells were transfected with pGFP plasmid and transfection efficiency was measured 48 hpt by measuring GFP expression by flow cytometry (**Figure 2A-F**).

Cells were transfected at a VCD of 2 x 10^6 cells/mL (considered low cell density (LCD)). In the control group, which did not receive lipid supplementation, a transfection efficiency of 60.9% was achieved. This efficiency was enhanced to 72% following the ME. Lipid supplementation strategies, however, did not yield significant differences in efficiency (**Figure 2A**). using the maximum concentrations of both fatty acids and cholesterol resulted in a modest increase in VCD at harvest, from 2.1 to 3.6 x 10^6 cells/mL (**Figure 2C**). These optimal concentrations were subsequently applied to the TT rAAV production at LCD, which maintained a similar viral titer to that of the control conditions (**Figure 2E**).

In contrast, when cells were transfected at a high cell density (HCD) of 6 x 10^6 cells/mL, a baseline transfection efficiency of 31% was observed, which increased to 59.4% with the ME prior to transfection. Notably, lipid supplementation strategies led to a marked improvement in transfection efficiency, achieving 50% with the minimum cholesterol concentration (**Figure 2B**). Across all experimental conditions, lipid addition also demonstrated a protective effect on VCD, obtaining higher cells at harvest (**Figure 2D**). The most successful lipid supplementation condition, using 5 pg/cell of cholesterol, was subsequently incorporated into the media during the TT process for rAAV

production in cells transfected at 6 x 10^6 cells/mL. This approach resulted in a doubling of transfection efficiency and a 233% increase in rAAV production compared to the baseline, which although low yield was obtained, it confirms the substantial benefits of the lipid supplementation protocol (**Figure 2F**).

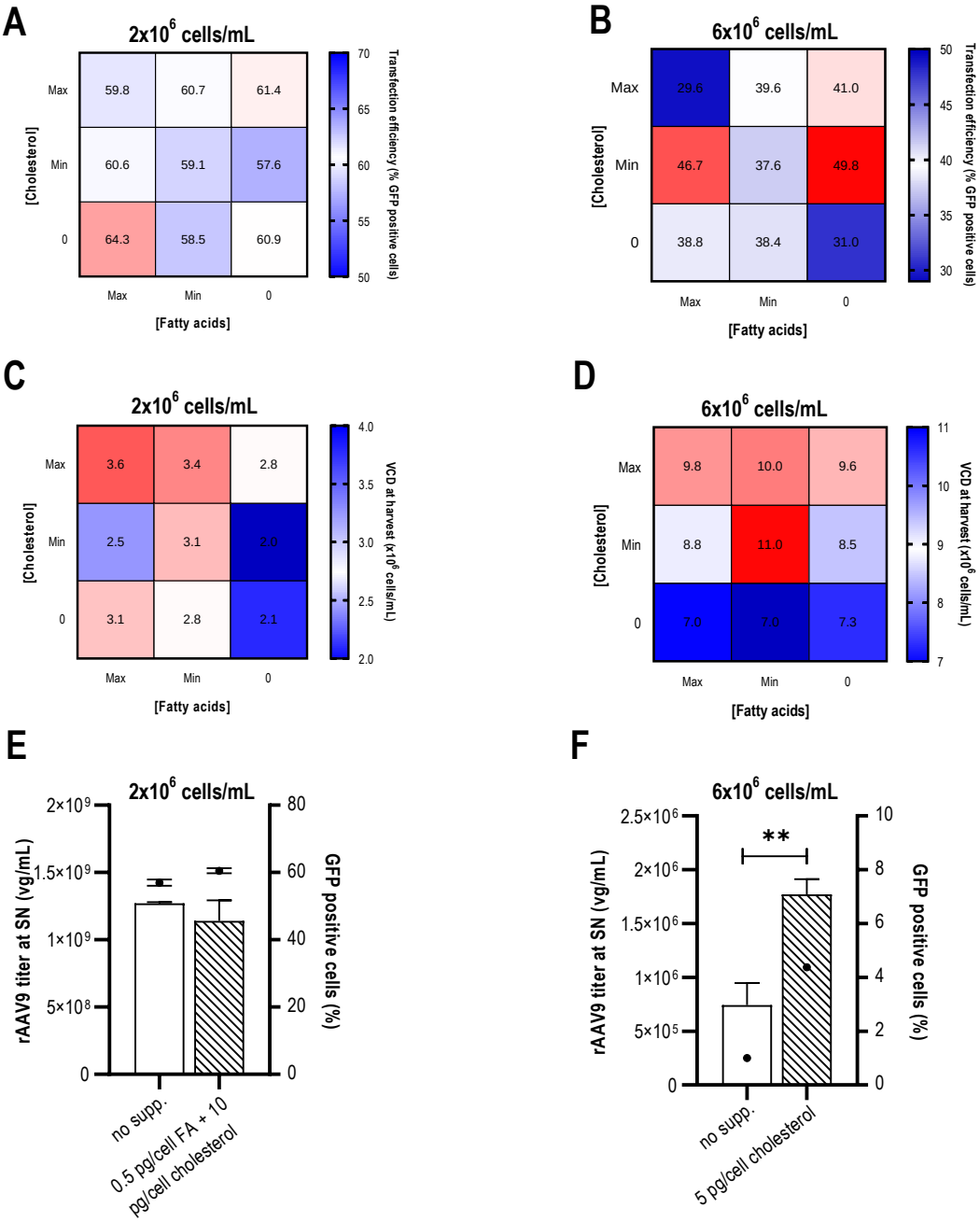


Figure 2. Analysis of cell transfection efficiency upon the addition of different lipid supplements to cell media 4 hbt (hours before transfection) at 2 and 6 x 10^6 cells/mL. Fatty acid was added at 0.5 pg/cell or 0.25 pg/cell for high and low concentration respectively. Cholesterol was added at 10 pg/cell and 5 pg/cell for high and low concentration respectively. **A and B.** Analysis of transfection efficiency by measure of GFP positive cells (%) 48 hpt upon lipid supplement addition at 2 x 10^6 cells/mL **(A)** and at and 6 x 10^6 cells/mL **(B).C and D**. Analysis of VCD at harvest upon lipid supplement addition at 2 x 10^6 cells/mL **(C)** and at and 6 x 10^6 cells/mL **(D). E and F.** Analysis of rAAV9 titer at cell supernatant of cells transfected at different cell densites with and without lipid supplementation at 2 x 10^6 cells/mL **(E)** and at and 6 x 10^6 cells/mL **(F).** Significance was calculated using an unpaired t-test. FA, fatty acids. VCD, viable cell density.

rAAV9 at the extracellular fraction in the EGE production method is associated with extracellular vesicles and presents higher potency compared to TT-derived rAAVs

Consistent with previous findings, the EGE method outperformed the TT method in terms of rAAV9 carrying GFP as GOI accumulation in the supernatant (SN) and a viral egress rate (71% in EGE compared to 58% in TT conditions) (**Figure 3A**). In this study, the association of rAAV9 harvested at cell supernatant with EVs was characterized using the EGE or TT protocols. To this end, rAAVs were collected via differential centrifugation, using a growth condition as a control. The supernatant fraction was obtained by centrifuging the cell culture at 1000 x g. To harvest the EVs and analyze their association with rAAVs, the supernatant was sequentially centrifuged as previously described for exo-rAAV harvest[154,345,348,354] and the pellet from a 100,000 x g centrifugation was collected.

First, DLS was used to analyze the particle size distribution based on their volume. In the cell supernatant fraction, the TT condition showed a bimodal distribution with particles around 20 nm, that may correspond to rAAVs, and a prominent peak of larger EVs around 340 nm. The EGE condition showed two peaks of reduced size compared to TT. In the growth condition, a single peak of 35-40 nm was observed (**Figure 3B**). Following the 100,000 x g ultracentrifugation, the particle size distribution in the EGE and the growth conditions showed a high similarity, while TT showed two distinct populations (**Figure 3C**).

Next, the mean particle size and concentration of the ultracentrifugation preparations was assessed using NTA. The analysis revealed a remarkably higher particle concentration in the EGE condition, which was comparable to that observed in the growth condition, and significantly higher than in the TT method (**Figure 3D**). Moreover, the EV particles collected from the EGE condition were smaller than those from both the TT and growth conditions (**Figure 3E**).

Finally, the potency of the rAAV in the various collected fractions was evaluated by assessing the transduction efficiency of different rAAV9 preparations on HEK293 cells, with wt AdV2 co-infection for signal amplification[352]. Interestingly, the analysis revealed that rAAVs derived from the EGE-exhibited greater potency than those from TT production method when evaluated directly from the supernatant fraction. In addition, post-ultracentrifugation, rAAVs from both EGE and TT methods exhibited similar potency (**Figure 3F**). Notably, those rAAV harvested after ultracentrifugation showed similar potency to those from the EGE supernatant fraction. This observation underscores the impact of enriching EV content on rAAV potency, potentially due to enhanced exo-rAAV recovery.

Moreover, the EGE method demonstrated a superior recovery efficiency of rAAV particles after ultracentrifugation, with 23% of vg recovered compared to 13% using the TT method. This suggests a higher association of rAAVs with EV particles in the EGE supernatant compared to TT, leading to more efficient recovery. This finding emphasizes the effectiveness of the EGE method in enhancing rAAV yield and potency through improved EV association.

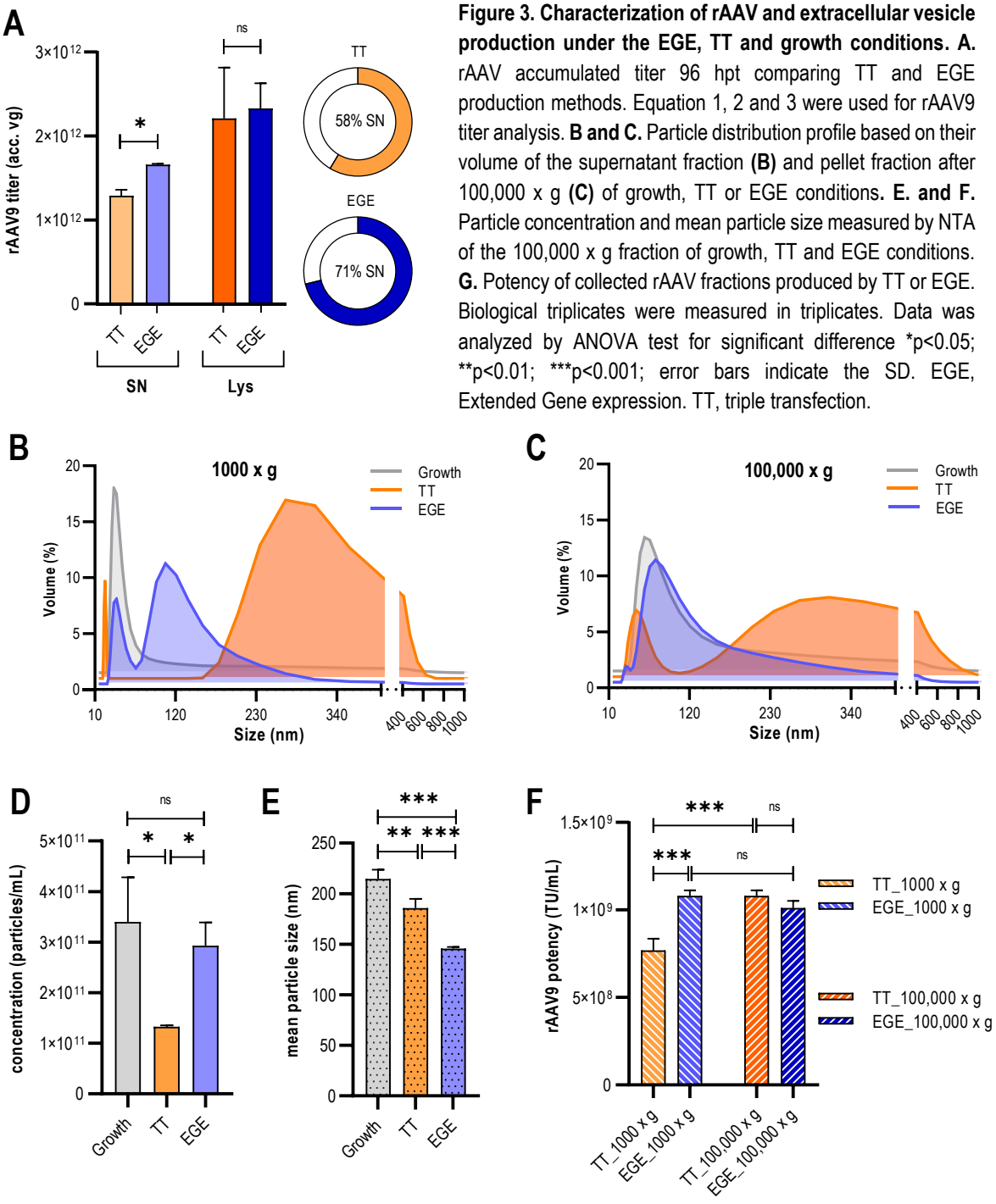


Figure 3. Characterization of rAAV and extracellular vesicle production under the EGE, TT and growth conditions. A. rAAV accumulated titer 96 hpt comparing TT and EGE production methods. Equation 1, 2 and 3 were used for rAAV9 titer analysis. **B and C.** Particle distribution profile based on their volume of the supernatant fraction **(B)** and pellet fraction after 100,000 x g **(C)** of growth, TT or EGE conditions**. E. and F.** Particle concentration and mean particle size measured by NTA of the 100,000 x g fraction of growth, TT and EGE conditions. **G.** Potency of collected rAAV fractions produced by TT or EGE. Biological triplicates were measured in triplicates. Data was analyzed by ANOVA test for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. EGE, Extended Gene expression. TT, triple transfection.

Metabolic engineering modulates extracellular vesicle production in HEK293 and CHO-S cells

Following the observation that the EGE method facilitates a greater association of rAAV9 with EVs compared to the TT method, potential modulators of the EV pathway were assessed. Specifically, the proteomic study detailed in **Chapter one** identified proteins overexpressed in EGE relative to the TT method involved in the EV pathway, particularly related to the exosome pathway. These proteins were targeted as potential modulators of the EV export route, and consequently, of the secretion of exo-rAAVs in mammalian cells[151,152,345].

The cDNA sequences of the identified proteins targeting the EV pathway were individually cloned into plasmids with the CMV as strong constitutive promoter. To test its overexpression, HEK293 cells were transfected with each of these plasmids, including pCD63, pVPS37B, pSNAP47, pRab32 and pSMPD3. The production of EVs was analyzed by NTA. Of note, EV production was measured relative to pMock control condition as an increase in EV production has been already observed when cells are transfected using PEI[355,356].

The results indicated that the overexpression of VPS37B, SNAP47 and SMPD3 in HEK293 cells led to an increase in EV production levels in the supernatant. A lesser increase was observed with the addition of CD63 and Rab32 (**Figure 4A**). The mean average size and the modal size of the EV were also characterized in each condition, the overexpression of CD63, VPS37B, and SNAP47 resulted in a slight increase in the size of the EVs compared to the control, with CD63 showing the highest modal size increase (**Figure 4B, C**).

Furthermore, analyses were extended to CHO-S cells to evaluate the transferability of these findings to another mammalian cell line. In this context, the overexpression of CD63 and VPS37B led to an increase in EV levels. Remarkably, this increase was more pronounced in CHO-S cells compared to HEK293 cells, with a log2 fold change (log2FC) of 3 to 4 relative to pMock transfection. This translates to a substantial increase from approximately $3.5x10^9$ EV/mL in exosomes in the control condition to around $1.5x10^{10}$ EV/mL with CD63 overexpression **(Figure 4D)**. Additionally, the overexpression of CD63 in CHO-S cells resulted in the production of smaller EVs, indicating different cellular responses compared to those observed in HEK293 cells **(Figure 4E, F**). The markedly greater increase in EV levels observed in CHO-S cells, compared to HEK293 cells, can likely be attributed to the initially lower baseline concentration of EVs in CHO-S cells. Specifically, HEK293 cells have a baseline EV concentration when transfected with pMock ten times higher than that of CHO-S cells, which inherently limits the potential for noticeable increases in HEK293 cells.

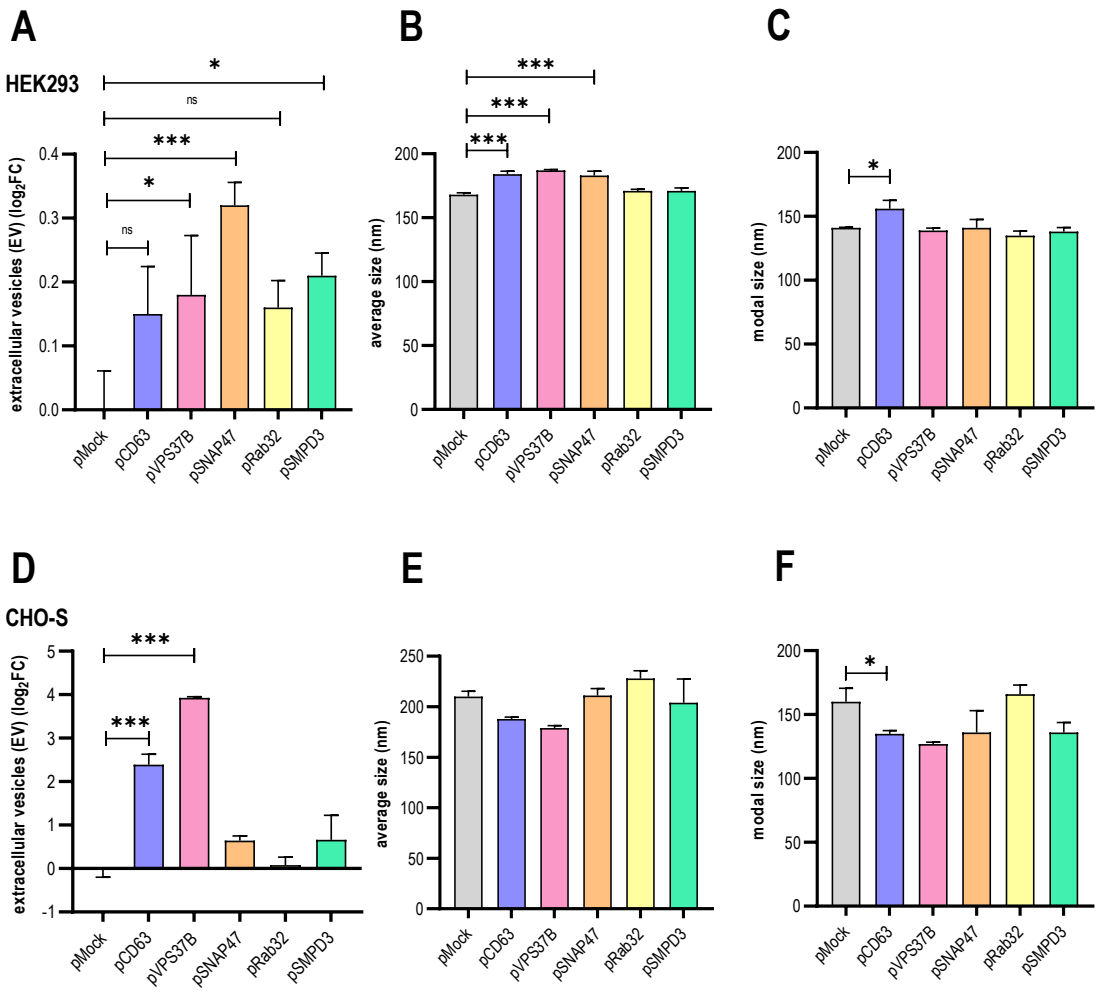


Figure 4. Characterization of extracellular vesicles (EV) in supernatants from different cell lines (HEK293 and CHO-S) following overexpression of candidate genes. Analysis of EVs using nanoparticle tracking analysis (NTA) comparing pCD63, pVPS37B, pSNAP47, pRab32 and pSMPD3 transfection to pMock. **A-C.** Results using HEK293 cell line. **D-F.** Results using CHO-S cell line. **A** and **D.** Log2 fold change (log2FC) in EV concentration across various gene overexpression conditions compared to the pMock transfection. **B.** and **E.** Average size of EVs produced under the different conditions. **C.** and **E.** Modal size of EVs produced under the different conditions. Biological triplicates were measured in triplicates. Data was analyzed by ANOVA test for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD.

Targeting of exo-rAAV pathway by metabolic engineering

Previously, the identified protein candidates, including pCD63, pVPS37B, pSNAP47, pRab32 and pSMPD3 showed consistent modulation of EV pathway in HEK293 and CHO-S cell lines. Next, its interaction with exo-rAAV pathway was explored.

Initially, the dependence on ESCRT-independent and ESCRT-dependent pathway for MVB loading and exosome release was examined to assess its influence on exo-rAAV production. The interaction with the ESCRT-I pathway was evaluated using two plasmids encoding a shRNA targeting the VPS37B gene (pshVPS37B1 and 2) as VPS37B depletion has been reported to lead to ESCRT-I destabilization[357]. Similarly, the ESCRT-I independent pathway was evaluated by shRNA targeting CD63 (pshCD63)[358]. The shRNA plasmids were co-transfected with TT plasmids and the rAAV9 production and its EV association was characterized.

Upon pshVPS37B.1 and 2, and pshCD63 expression, a reduction in EV size profile measured by DLS remarkably shifted to a narrower peak at 25 nm compared to TT particle size distribution (**Figure 5A**). In addition, rAAV9 titer in the supernatant decreased when using the shRNAs compared to standard TT condition and to a scrambled shRNA control (pshCNT) (**Figure 5B**).

Next, the previously identified plasmid candidates affecting the EV production were co-transfected with TT plasmids to analyze its effect on exo-rAAV production (**Figure 5C, D**). Upon CD63 overexpression, particle size distribution profile was highly comparable to the TT control production, with a primary peak at 50-60 nm and a secondary peak at 400 nm, while VPS37B overexpression led to a reduction in EV size (**Figure 5C**). Of note, both CD63 and VPS37B increased extracellular rAAV titers (**Figure 5D**).

Upon overexpression of SMPD3 which breaks down sphingomyelin to ceramide, extracellular rAAV titers increased (**Figure 5D**) and DLS particle size distribution resembled the profile of the TT production (**Figure 5C**). To test the effect of ceramide, it was added as a supplement to a TT production and to the condition overexpressing SMPD3. Both conditions significantly decreased the extracellular rAAV titers compared to the TT production (**Figure 5D**).

To evaluate the next step in exosome generation, Rab32 and SNAP47, involved in the fusion of the MVB to the cell membrane were individually overexpressed. Co-transfection with pSNAP47 resulted in a particle distribution characterized by a single, more prominent peak in the total volume percentage, as well as an increase in rAAV titer in the supernatant (**Figure 5C, D**). Contrarily, Rab32 overexpression showed no increase in rAAV, associated with a decrease in particle size (**Figure 5C, D**).

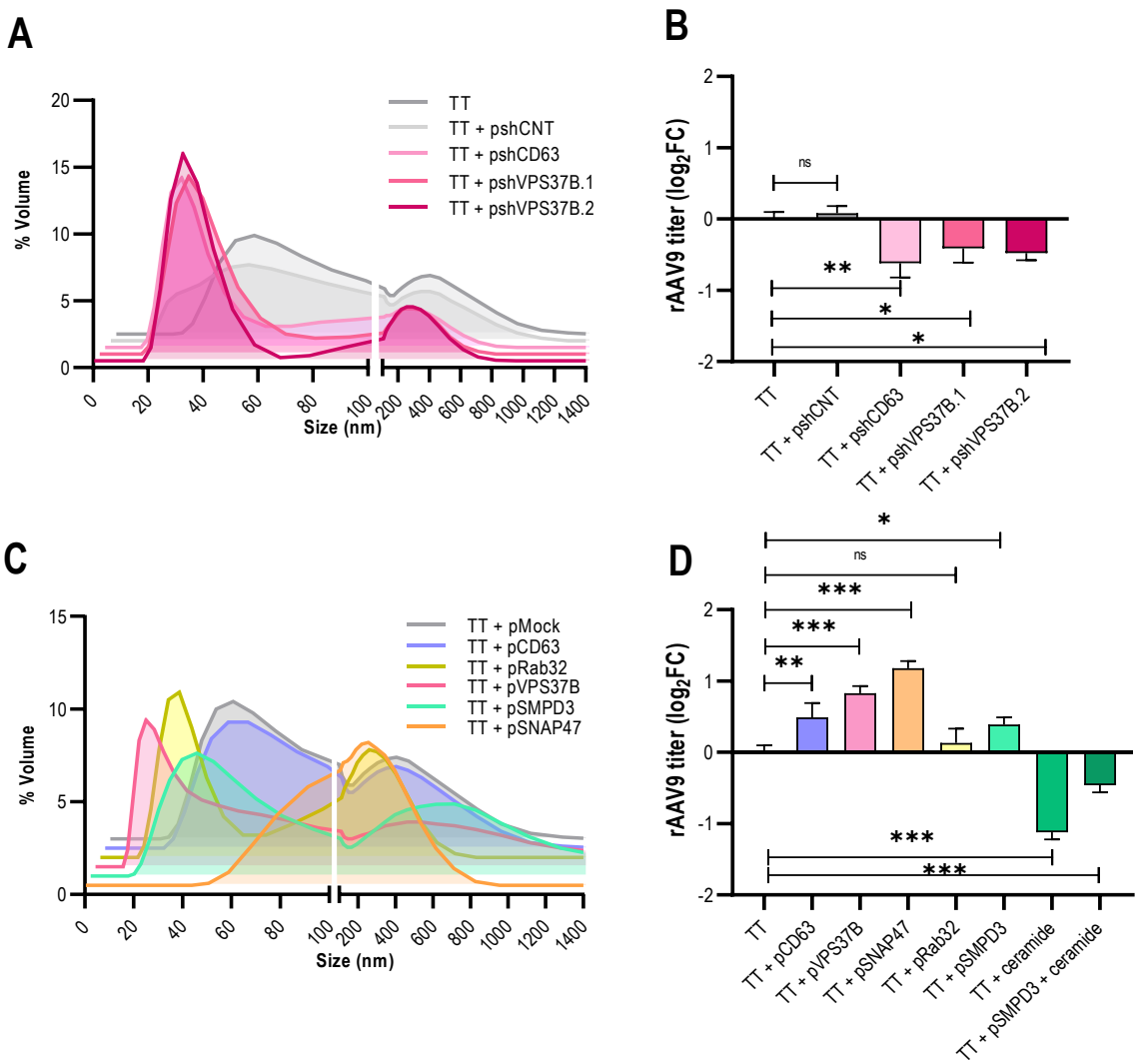


Figure 5. Characterization of rAAV9 titer and vesicle size distribution profile in cell supernatant fraction. A. Comparison of DLS plot between the different shRNA generated in co-transfection with TT plasmids. **B.** rAAV9 titer expressed as log_2FC (fold change) of standard triple transfection plasmids (TT) adding a different test plasmid harboring a shRNA targeting different genes in each condition compared to standard triple transfection with pMock**. C.** DLS plot of TT adding different tested plasmids, showing different volume of extracellular vesicles at different size at cell supernatant**. D.** rAAV9 titer expressed as log_2FC (fold change) of standard triple transfection plasmids (TT) adding a different test plasmid overexpressing different genes in each condition compared to standard triple transfection with pMock. For DLS plot, average of three different measures is shown. For rAAV9 titer, biological triplicates were measured in triplicates. Data was analyzed by ANOVA test for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. TT, triple transfection.

Metabolic engineering increases extracellular rAAV production and viral egress rate

Following the observations that rAAV titers in the cell supernatant fraction were influenced by the identified proteins targeting the EV pathway, further experiments were conducted to assess the impact of protein combinations on rAAV9 secretion. Specifically, the effects of co-expressing two proteins and rAAV9 titer in the cell supernatant fraction were characterized (**Figure 6A-D**).

As a result, specific combinations increased rAAV secretion. The pairing of CD63 with SMPD3 resulted in the most significant increase in rAAV titer compared to the TT condition. Additionally, the combinations of SNAP47 with SMPD3, and SNAP47 with VPS37B, also enhanced extracellular rAAV titers, although to a lesser extend (**Figure 6B**).

Furthermore, rAAV egress under these conditions was evaluated (**Figure 6C, D**). When comparing individual expression of each candidate protein during TT, CD63 showed the highest percentage of rAAV9 in the cell supernatant fraction (**Figure 6C)**. Among the protein combinations, CD63-SMPD3, SNAP47-VPS37B, and SNAP47-SMPD3 showed the most pronounced rAAV egress rates (**Figure 6D)**.

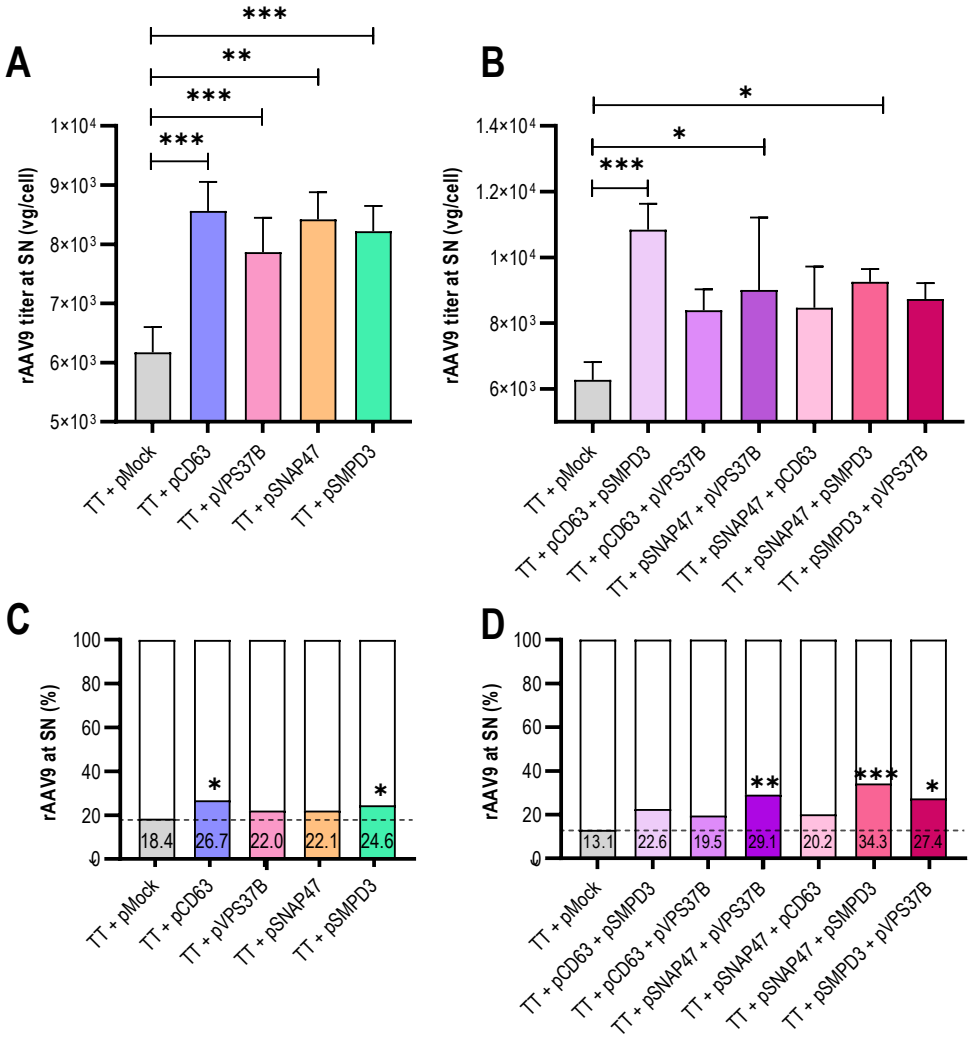


Figure 6. Characterization of rAAV9 titer at extracellular media and produced exo-rAAVs. A. rAAV9 titer of quadruple transfection using standard triple transfection plasmids adding one of the candidate genes. **B.** rAAV9 titer of triple transfection plasmids co-transfected with two candidate genes**. C. and D.** Biological triplicates were measured in triplicates. Data was analyzed by ANOVA test for significant difference *p<0.05; **p<0.01; ***p<0.001.

Extracellular rAAV increase is correlated with higher extracellular vesicle secretion

To establish a correlation between rAAV secretion and exosome release, dot blot analyses were performed using CD63, a known exosome marker. Enhanced levels of the exosome marker were detected in all conditions that involved overexpression of CD63, either alone or in combination with other proteins (**Figure 7A, B** respectively). In addition, higher exosome secretion levels compared to the control condition were observed for VPS37B overexpression and SMPD3 and VPS37B combination (**Figure 7A, B**).

Further, the supernatant from the CD63-overexpressing samples was examined using cryo-Transmission Electron Microscopy (cryo-TEM). The images revealed exosome particles encapsulating rAAVs, providing visual confirmation of the exosome-mediated rAAV shielding (**Figure 7C, D**). This data supports the hypothesis that enhancing exosome production through specific genetic modifications not only increases the secretion of exosomes but also correlates with higher levels of extracellular rAAV.

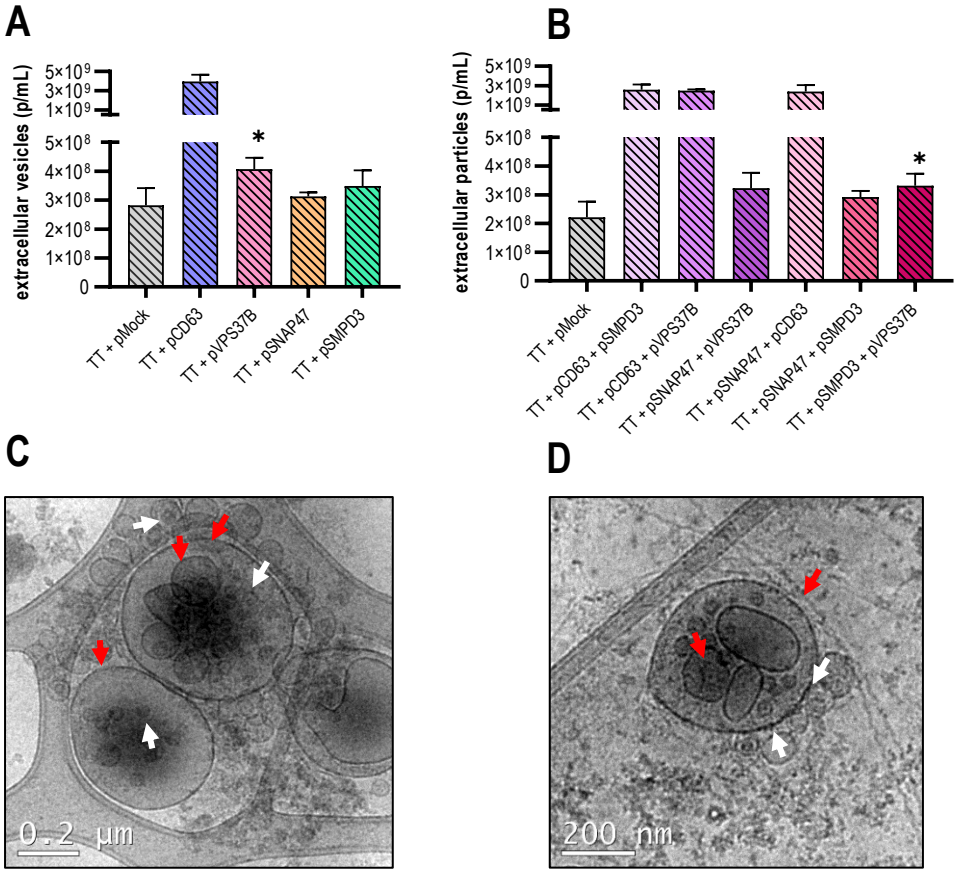


Figure 7. Characterization of rAAV9 titer at extracellular media and produced exo-rAAVs. A. and B. Extracellular vesicle concentration measured by dot blot correlation to CD63 antibody. Results are shown as relative intensity measured with ImageJ compared to a standard curve of known extracellular vesicle concentration. Data was analyzed by t-test for significant difference, comparing each condition to TT+pMock control *p<0.05; **p<0.01; ***p<0.001. **C. and D.** Cryo-transmission electron microscopy of harvested supernatant of transient transfection with CD63 overexpression. The image shows rAAV particles associated to exosomes in supernatant fraction at harvest. White arrows indicate rAAV particles and red arrows indicate exosome membranes.

Determination of optimal Metabolic Engineering combination for extracellular rAAV harvest

Having observed that pSNAP47, pSMPD3, and pVPS37B contributed positively to rAAV yield in the supernatant, efforts were directed towards identifying the optimal combination of these genes to establish a HEK293 cell line capable of maximizing the rAAV egress rate.

A DoE, specifically a simplex design, was selected for this investigation. This approach allows for the determination of not only the relevant genes but also their optimal ratio combination for each desired response: Esn, corresponding to the egress rate and Psp, measuring the specific production. The experimental design matrix, response and statistical analysis are shown at **Table 3**. Data were fitted to a quadratic second-order polynomial equation (**Equation 6**) using a linear regression analysis. The model coefficients were determined using Type III partial sum of squares, which allows for isolation of each variable's effect in the presence of other variables (**Table 4**).

Table 3. Matrix design, results, and ANOVA analyses for Simplex design for optimization of rAAV9 production in HEK293 cell line.

	Variables			Response	
Experimental Run	**pSNAP47 (%)**	**pSMPD3 (%)**	**pVPS37B (%)**	**Esn (%)**	**Psp (vg/cell)**
1	0	0.5	0.5	51.68	3575.82
2	0.16	0.16	0.66	29.40	4903.38
3	0.5	0.5	0	17.13	11745.95
4	0	1	0	18.14	13865.20
5	0	0	1	40.58	3970.19
6	0.16	0.16	0.16	37.84	4024.09
7	1	0	0	22.30	7938.26
8	0.5	0	0.5	21.79	7666.83
9	0.33	0.33	0.33	21.06	9600.76
10	0	1	0	20.48	11800.18
11	0	0	1	28.21	7196.95
12	0.5	0.5	0	16.55	13150.01
13	0.16	0.16	0.16	24.23	8033.15
14	1	0	0	30.92	5398.60

Model	**Multiple R^2**	***p* value** [(a)]	**Lack of fit test, *p* value** [(b)]
Esn (%)	0.7992	0.0113	0.2183
Psp (vg/cell)	0.7421	0.0283	0.1898

	Esn (%)		**Psp (vg/cell)**	
Parameters	**Sum of squares** [(c)]	***p* value**	**Sum of squares**	***p* value**
Model	1035.97		1.192×10^8	
Linear Mixture	398.39	0.0244	6.653×10^8	0.0217
AB: pSNAP47 + pSMPD3	64.37	0.1972	8.297×10^8	0.2412
AC: pSNAP47 + pVPS37B	169.73	0.0517	2.982×10^8	0.4697
BC: pSMPD3 + pVPS37B	456.64	0.0057	4.339×10^8	0.0201

[(a)] *p values* under 0.05 are considered statistically significant with 95 % confidence, and under 0.1, statistically significant with 90 %.

[(b)] p values associated to lack of fit test above 0.05 mean that the hypothesis arguing that the model is suitable cannot be rejected.

[(c)] Type I sum of squares was used for lineal mixture calculation.

Regression analysis for both responses (Esn and Psp) showed that the model fitted the data with a coefficient R^2 of 0.7992 and 0.7421. respectively. Further, the statistical significance of the model was confirmed by ANOVA analysis. The Fisher's F-test associated *p* value of 0.0113 and 0.0283 indicated that both models were significant. The lack of fit parameter was non-significant (*p* value 0.2183 and 0.1898). Values of the linear mixture were significant, while combination of SMPD3 and VPS37B were the only significant combinations in both models (see **Table 3**).

The simplex design responses were used to calculate the optimal plasmid ratios for rAAV9 production (**Figure 8**). Contour plots were constructed using the model to visually represent how different combinations of plasmids affect each specific response. The maximal egress of rAAV9 (Esn) was achieved when SMPD3 and VPS37B ware combined at ratios of 0.415 and 0.585, respectively, achieving a 50% egress rate (**Figure 8A**). The specific production of rAAV9 in the cell lysate (Psp) was maximal when SMPD3 and SNAP47 were combined at ratios of 0.79 and 0.21, respectively, achieving 1.26 x 10^4 vg/cell. The inclusion of VPS37B had minimal influence on Psp (**Figure 8B**).

When both responses were maximized to obtain both optimal Esn and Psp, the combination of SMPD3 and VPS37 at 0.877 and 0.123 ratio led to a 31.6% of viral particles at supernatant and 8631.65 vg/cell of specific production at cell lysate fraction.

In parallel, a control condition was employed where pMock was added at a concentration of 9x10^{10} copies/mL instead of the tested plasmid combination. The viral egress rate in this condition was 19.7%, and the Psp measured at 5922.46 vg/cell. Thereby, demonstrating that the optimal plasmid combinations derived from the simplex design response outperformed the control condition in terms of both egress and specific production.

Table 4. Coefficients in Terms of Coded Factors of mixed factorial design simplex in Table 3.

	Esn (% SN)	**Psp (vg/cell)**
Parameters	**Coefficient Estimate**	**Coefficient Estimate**
A. pSNAP47	27.49	6737.61
B. pSMPD3	20.79	12230.43
C. pVPS37B	35.12	5708.34
AB: pSNAP47 + pSMPD3	-26.34	9458.03
AC: pSNAP47 + pVPS37B	-51.28	6796.04
BC: pSMPD3 + pVPS37B	84.10	-25925.21

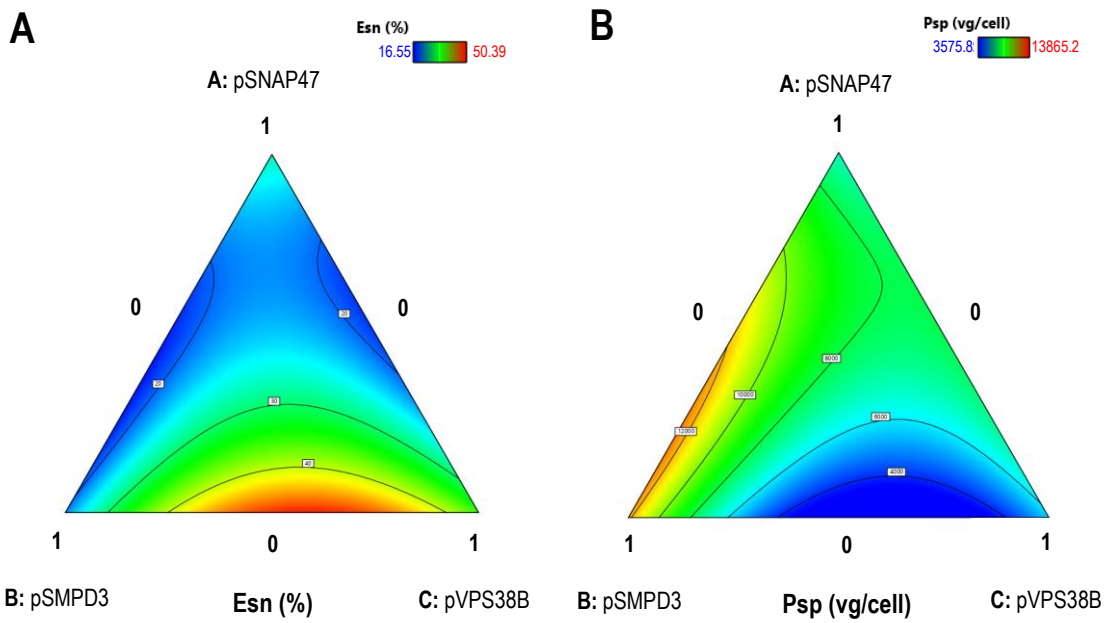


Figure 8. Simplex analysis of rAAV9 production. A. Contour plot for rAAV9 egress at supernatant (Esn) measured as the percentage of rAAV9 detected at cell supernatant. **B.** Contour plot for specific production response (Psp) measured as total rAAV9 specific production per cell at cell lysate fraction, involving cell pellet and cell supernatant (vg/cell). Each axis of the triangle represents one of the three plasmids (A: pSNAP47, B: sPMPD3, C: pVPS37B) at varying concentrations, demonstrating their proportional effects on each response. The color gradient and the contour lines provide a visual representation of response variable levels, with warmer colors indicating higher values in each response.

DISCUSSION

The EGE production method outperforms the standard TT for rAAV manufacturing[214], primarily due to two main factors: first, the EGE method exploits the phenomenon of rAAV release into cell culture media and second, the RT step extends the production timeframe, thereby leading to higher accumulated rAAV particles at the cell supernatant[295,297].

The obtained results demonstrate that the EGE method not only increases the rAAV release into the cell supernatant compared to the TT but also enhances EV co-production. Moreover, the enhanced transduction capacity of the EGE-derived supernatant fraction, in comparison to that harvested from TT, suggests that rAAVs are shielded by EV particles contributing to its superior transduction properties as previously reported[152,346–348]. Therefore, it suggests a potential link between the EGE operation mode and higher exo-rAAV production.

Building on this findings, it is proposed that viral egress mechanism might be associated with EV release, with rAAV particles secreted from the cytoplasm via exosomes[156]. Such association aligns with previous findings based on the novel AAV membrane-associated accessory protein (MAAP) which has been shown to be a viral egress factor by promoting rAAV exit from the cells by interacting with the outer surface of extracellular vesicles[89]. Therefore, to exploit the rAAV release property, putative modulators of the EV egress pathway potentially influencing the rAAV egress were identified in the proteomic study presented in **Chapter one**. This aims to design more efficient production strategies that focus on harvesting the extracellular fraction only. Specifically, key metabolic targets within the molecular pathways of MVB loading and exosome release, particularly those overexpressed in the EGE method compared to TT, were identified. This candidate proteins, including SNAP47, Rab32, VPS37B, SMPD3 and CD63, proved effective as modulators of the exosome pathway in mammalian cells, such as HEK293 and CHO-S.

To explore the association of the exosome pathway with rAAV release, various stages including the maturation of early endosomes into MVBs, the loading of rAAVs into these MVBs (VPS37B, CD63, SMPD4), and their subsequent fusion with the cell membrane for release were targeted (SNAP47 and Rab32), as illustrated in **Figure 1,** showing putative rAAV egress pathway. This identified proteins when overexpressed during TT for rAAV production led an increase in extracellular rAAV9 titer and egress rate at various levels. The increase of rAAV titer at extracellular fraction together with differential EV profile evidenced the effective modulation of exo-rAAV secretion pathway, opening new avenues for enhancing rAAV production efficiency at supernatant.

To further understand the rAAV release, shRNAs targeting proteins involved in ESCRT-dependent (VPS37B) and independent (CD63) pathways were used. The reduction in rAAV titer in the extracellular fraction was observed in both pathways, thereby suggesting that rAAV release may be associated to different mechanisms. In future studies, the design of shRNA against other targets of the EV pathway and may help identify essential effectors of rAAV egress[359]. Further, the microvesicle secretion pathway, which refers to the vesicles that shed directly from the cell's plasma membrane (depicted in **Figure 1**) may also be considered to generate a wider picture of rAAV release origin[360]. Combination of siRNA targets may help explain further the rAAV release mechanisms.

Interestingly, metabolic engineering through the expression of CD63, not only increased rAAV production in the cell supernatant, but also significantly increased co-production of EVs, thereby exo-rAAVs. It is hypothesized that targeting CD63 to enhance exo-rAAV9 production could yield more significant outcomes when a CD63-overexpressing stable cell line is generated. Such a cell line, when used for rAAV production by TT, could lead to substantial benefit in rAAV manufacturing. This hypothesis is further supported by the findings of Schiller *et al*., who developed a HEK293 cell line overexpressing the CD9 tetraspanin[153]. While their approach increased exo-rAAV production, it did not concurrently increase rAAV titers in the supernatant compared to the control cell line. The proposed strategy with CD63, on the other hand, appears to address both aspects effectively. The potential applications of a CD63-overexpressing stable cell line may be of use when higher exo-rAAV yields are required, allowing to explore their beneficial properties and future applications compared to standard rAAVs. For example, the engineering of ligands and transmembrane receptors onto the surface of exo-rAAVs could be used to expand transduction specificity of tissues[154]. In addition, scalable purification methods such as chromatography, also used for viral particle purification should be considered, as extracellular purification using ultracentrifuge methods is not suited for exo-rAAV yield recovery at large scale[361].

In this research, the overexpression SNAP47, VPS37B, and SMPD3, while not substantially increased EV co-production, significantly boosted viral egress of rAAV9. Using a DoE strategy the optimal combination of proteins resulting in the highest viral egress rate was identified, with the co-expression of VPS37B and SMPD3 being the most effective in maximizing rAAV release into the supernatant fraction reaching near 50%. This discovery holds significant value for the future development of novel HEK293 stable cell lines with improved metabolic engineering characteristics. The generation of a HEK293 stable cell line harboring high copy numbers of *smpd3*

and *vps37b* genes stably integrated could be beneficial to enhance extracellular rAAV production improving yield and harvest at cell supernatant fraction by the TT production method.

Considering the advancements into the generation of a mammalian cell line with improved rAAV egress properties, theoretically, increasing the VCD at the time of transfection could proportionally lead to a higher volumetric productivity of viral particles. However, a notable reduction in transfection efficiency and productivity is observed at high cell densities, known as the CDE[190,281,287], imposing a significant bottleneck for TGE manufacturing approach[286]. In this context, the CDE was addressed using the proteomic approach results described in **Chapter one**. Cells cultured at HCD engage in energy conservation strategies by reducing *de novo* lipid biosynthetic pathways, as also reported by Lavado *et al*. under similar growth conditions[316]. This study demonstrates that lipid supplementation, particularly the addition of synthetic cholesterol to high cell density culture media, can partially counteract the CDE. This protocol showed comparable efficacy to when ME is performed before transfection, resulting into higher rAAV production. These findings suggest that cell membrane lipid composition might be related to the capability of the cell of internalizing DNA/PEI complexes[316]. Importantly, the lipid supplementation strategy to cell culture media explored here is simple yet markedly effective in facilitating efficient transfection at high cell density, potentially facilitating the scale-up of biopharmaceutical products using the transient transfection approach.

In summary, this research successfully identifies key targets for enhancing exo-rAAV production and viral egress rates, underscoring the benefits of creating stable cell lines that can modulate host cell metabolism to improve the rAAV production process. Such stable cell lines with enhanced egress rate of rAAV could benefit from the strategies to improve transient transfection at HCD, such as with the lipid supplementation feeds reported here. The combination of these strategies could substantially improve rAAV production processes, by improving rAAV yield and decreasing production costs.

CONCLUSIONS

This study evaluates the Extended Gene Expression (EGE) method and its enhancement of rAAV9 production compared to the standard Transient Triple transfection (TT) approach. The findings indicate that the EGE method increases the association of rAAV9 with extracellular vesicles (EVs), potentially due to increased viral egress mechanisms and enhanced potency. Proteins upregulated in the EGE method, identified through a proteomic study focusing on the EV pathway, were selected for metabolic engineering modifications. This approach has significantly increased EV production in HEK293 cells and demonstrated comparable success in CHO-S cells. Additionally, by overexpressing EV-targeting proteins, the exo-rAAV pathway was specifically enhanced, leading to an increased rate of rAAV egress and co-produced extracellular vesicles.

Moreover, using the Design of Experiments (DoE) approach, optimal gene combinations, particularly SMPD3 and VPS37B, were identified to maximize the yield of extracellular rAAV. The study also proposes media supplementation strategies that improve transfection efficiency at High Cell Densities (HCD), correlating with higher extracellular rAAV yields and enabling a more scalable implementation of the TT method for large-scale production.

Collectively, these findings provide valuable insights into effective molecular modulators that enhance the TT method for rAAV production at different levels. The generation of stable cell lines incorporating the metabolic engineering modifications lays the groundwork for future strategies aimed at more efficient production and export of rAAV in the cell supernatant fraction. This approach holds promise to reduce production costs and facilitate downstream process, thereby contributing to enhance the scalability and affordability of rAAV-based products.

CHAPTER THREE

Strategies to circumvent endogenous *E1A* gene expression bottleneck in HEK293 cells for enhanced rAAV production

The contents of this Chapter generated the patent entitled: *Method for production of gene therapy vehicles in mammalian cells*, EP24159714.5.

INTRODUCTION

The HEK293 cell line is currently the workhorse for recombinant adeno-associated virus (rAAV) production using the Transient Gene Expression (TGE) system via the triple transfection (TT) method[210,296,362]. These cells were originally transformed with the *E1* gene by transfection with sheared adenovirus 5 (AdV5) genomic DNA, resulting in the constitutive expression of E1A and E1B proteins[181,182]. The E1A protein plays a critical role in initiating rAAV replication, by enhancing transcription from AAV *rep* gene promoters p5 and p19[47], and transactivating early AdV promoters[105]. Additionally, E1A drives host cells into the S-phase of the cell cycle, thereby modulating the host cell environment to promote viral replication[110].

The inherent expression of *E1* AdV gene in HEK293 cells renders them particularly suited for rAAV production. The complete set of AdV helper genes are provided through a plasmid containing *E2A, E4 and VA RNAs* lacking the *E1* gene, as it is already present in HEK293 cell line[99,101,106,181,182]. In another plasmid, *rep* and *cap* genes for AAV replication and encapsidation are provided. Finally, the gene of interest is supplied in a third plasmid, flanked by Inverted Terminal Repeats (ITRs).

Another characteristic of HEK293 cells is that they are adapted to serum-free suspension conditions, enhancing the scalability potential of rAAV production[210,296,362]. The rAAV production by the TGE approach by TT, not only enables the production of substantial viral particle yields but also offers flexibility in product development. However, the high amount of plasmids required in the TT adds cost and variability to the process and leads to a system limited in time as plasmid DNA does not integrate into the host cell's genome. Thus, new methods have been developed to enable robust rAAV production at large-scale such as the stable gene expression (SGE) by the generation of stable producer cell lines[246]. In the context of rAAV production, the development of packaging/producer stable cell line platforms has been under research for decades as a valuable method for biopharmaceutical manufacturing[246]. This approach involves integrating the necessary genes for rAAV production into the host cell's genome. However, when HEK293 are used for stable rAAV producer cell line development, inducible promoters are needed to address the toxicity, induced by viral protein expression, specifically of *rep* and AdV helper genes due to *E1A* transactivation[58,197,198]. Although HEK293-based stable producer cell lines have been generated, they exhibit low titers compared to the TGE method[275]. Moreover, the process of generating a stable cell line is laborious and time-consuming and a new cell line is required for each rAAV product, being much less flexible compared to the TGE system.

In an effort to combine the advantages of the TGE for rAAV production with an extension of the culture time to increase the total rAAV yield, a novel approach named extended gene expression (EGE) was developed[294]. This production method was originally devised for HIV-1 virus-like particles (VLP), successfully expanding the production timeframe from 72 hpt to 240 and significantly increasing recombinant protein production levels[295]. Additionally, the applicability of the EGE for enhancing rAAV titers has also been demonstrated[214]. The core strategy of the EGE method involves regular medium exchanges to remove waste by-products and refresh culture medium, combined with periodic re-transfections of the cell culture to increase the availability of plasmid DNA to extend the production period[214]. Importantly, the transferability of this method for continuous biomanufacturing processes based on perfusion has been demonstrated, highlighting its potential for large-scale rAAV production[295,297].

As presented in **Chapter one**, the application of the EGE method significantly enhanced both the specific production and the viral egress rate of rAAV9 compared to the standard TT approach. A multiplexed proteomic methodology was used to examine the changes in host cell responses attributable to the EGE method, comparing: a non-transfected HEK293 cell growth condition serving as a control, rAAV production by the standard TT, and rAAV production by the EGE protocol. This comparative analysis provided a comprehensive perspective on the molecular mechanisms governing rAAV production by TT and the EGE methods regarding viral and host cell protein dynamics during rAAV production. Changes in host cell proteins during rAAV production were extensively examined in **Chapter one** and exploited in **Chapter two** to improve rAAV production methods. Here, the focus is shift to adenoviral protein dynamics, specifically the expression of E1 proteins over time in rAAV production conditions.

MATERIALS AND METHODS

Mammalian cell line and culture conditions

The cell line used in this work is a serum-free suspension-adapted HEK293 cell line (HEK293SF-3F6) kindly provided by Dr. Amine Kamen from the National Research Council of Canada (NRC, Montreal, Canada). Cells were cultured in disposable polycarbonate 125 mL vented shake flasks (Corning®, Tewksbury, NY, USA) at 37ºC, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland).

Cell culture media was HyCell™ TransFx-H (GE Healthcare, Chicago, IL, USA) supplemented with 4 mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA) and 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, ThermoFisher Scientific, San Jose, CA, USA). Cell concentration and viability were determined using the NucleoCounter® NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions.

Plasmid design and generation

Plasmids used for triple transfection were kindly provided by Dr. Fatima Bosch from Center for Animal Biotechnology and Gene Therapy (CBATEG), at Universitat Autònoma de Barcelona (UAB, Bellaterra, Spain). Specifically, the plasmids pHelper, pGG2_CMV_GFP (pGFP_ITR, hereafter) and pRep2CapX were used. The expression cassette is flanked by two inverted terminal repeats (ITRs) of AAV2. The pHelper plasmid codifies for the adenovirus helper genes, namely: *E2A, E4* and *VA RNA I* and *VA RNA II.* Plasmid pGFP_ITR, encodes for enhanced GFP under the control of the cytomegalovirus (CMV) promoter. The plasmid pRep2CapX encodes the Rep proteins from AAV2, while the "X" represents the resulting rAAV serotype given by the capsid gene sequence. Here, pRepCap1, pRepCap6, pRepCap8 and pRepCap9 were employed.

To construct the pE1_cnt plasmid, the *E1A* and *E1B* gene fragment from AdV5 genome (AC_000008.1) nucleotides (nt) 194 to 3509 were cloned into pcDNA3.1 + /C-(K)DYK plasmid backbone and bGh PolyA signal was inserted downstream (GenScript Europe) including endogenous promoter regions[204,257,363]. EcoRI enzyme restriction site was inserted into the 501 position for posterior sub-cloning.

pMock was used as transfection control plasmid, as it does not have any mammalian promoter or coding DNA sequence (CDS). It was constructed by the ligation of the pGag::eGFP backbone[350,351].

To check for individual elements contributing to cellular growth, viability and *E1* gene expression, elements in triple transfection plasmid were separated. pHelper plasmid was divided into two plasmids carrying AdV helper genes: pE2 and pVA_RNA_E4. The pVA_RNA_E4 plasmid contained nucleotides from 10435-11166 nt for VA RNA sequence and 32648-35835 nt for full E4 open reading frame (ORF) from AC_000008.1 (AdV5 complete genome). The pE2 plasmid was constructed containing 22188-27533 nt corresponding to E2 ORF.

As for pRep2Cap1, Rep and Cap ORF were separated. To construct pRep_wt plasmid, Rep ORF from serotype 2 was PCR amplified using Q5 High-Fidelity DNA polymerase (New England Biolabs (NEB), Ipswich, MA, USA) from pRep2Cap9 plasmid using phosphorylated primers shown in **Table 1** following manufacturer's instructions. Resulting fragments were purified and ligated using T4 DNA ligase (ThermoFisher Scientific, San Jose, CA, USA). To construct pCMV_CAP1 plasmid, VP1 ORF from pRep2Cap1 plasmid was amplified using Q5® High-Fidelity DNA polymerase (NEB, Ipswich, MA, USA) following manufacturer's instructions. Primers were designed to generate overlaps of 20 nt with homology to CMV fragment. CMV enhancer and promoter region was cloned from pGag_GFP plasmid. Two-fragment assembly was performed using the NEBuilder® HiFi DNA Assembly mix (NEB, Ipswich, MA, USA) following manufacturer's instructions with 50 ng of vector and a vector-insert ratio of 1:2. Primers used are shown in **Table 1.**

Plasmids were confirmed by Sanger sequencing (Servei de Genòmica, UAB, Bellaterra, Spain) and amplified in a chemicompetent *E. coli* DH5α strain produced *in house* growing in LB medium (Conda, Madrid, Spain) supplemented with 10 µg/ml kanamycin (Sigma-Aldrich, St. Louis, MO, USA) or 100 µg/ml ampicillin (Sigma-Aldrich, St. Louis, MO, USA) depending on the *E. coli* antibiotic resistance present on each plasmid. Plasmid purification was carried out using the Endofree Plasmid MIDI kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Quantity of isolated DNA was assessed using Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, San Jose, CA, USA).

Transient transfection for rAAV production

Transfections were carried out at a cell density of 2 x 10^6 cells/mL using an equimolar plasmid concentration for each of the triple transfection plasmids: pHelper, pRep2CapX, pGFP_ITR and an additional plasmid (either pE1_cnt or pMock) of 3 x 10^{10} copies/mL. Total DNA was normalized to 1 µg/ml using pMock when necessary. Transfections were carried out in 125 mL erlenmeyer flasks using 10 mL of total culture volume.

Table 1. List of primers used for plasmid generation and synthetic sequences. Sequences added to generate sticky ends or homologies for DNA assembly are underlined. Introduced mutations are shown in lowercase. Primers were designed avoiding possibilities of self-dimerization and primer-dimer formation using Primer3Plus.

Name	Objective	Sequence (5'-3')
Fwd_Rep2_wt	Amplify Rep78 ORF from pRep2Cap9 plasmid	TGTTTCCTTGCGTATTGGGC
Rev_rep2_wt		ACCTGGTTTAAGTCATTTATTGTTC
Fwd CMV-CAP1	Amplify Cap1 ORF from pRep2Cap1 plasmid	GTGAACCGTCAGATCCGCTAATGGCTGCC GATGGTTATCT
Rev CMV_CAP1		TATGTAACGCGGAACTCCATCTCACTGGCC GTCGTTTTAC
Fwd_CMVenh	Amplify CMV enhancer and promoter region from pGag_GFP plasmid	ATGGAGTTCCGCGTTACATA
Rev_CMVenh		TAGCGGATCTGACGGTTCAC
Fwd_GFP_qPCR	qPCR for rAAV titer determination	AGCTGGACGGCGACGTAA
Rev_GFP_qPCR		GTGCAGATGAACTTCAGGGTCA
GFP gBlock	gBlock used for standard curve in qPCR	AGCTGTTCACCGGGGTGGTGCCCATCCTG GTCGAGCTGGACGGCGACGTAAACGGCCA CAAGTTCAGCGTGTCCGGCGAGGGCGAGG GCGATGCCACCTACGGCAAGCTGACCCTG AAGTTCATCTGCACCACCGGCAAGCTGCCC GTGCCCTGGCCCACCCTCGTGAC

For transient transfection, 25-kDa linear polyethilenimine (PEI) (PEIPro®, Polyplus-transfection SA, Illkirch-Graffenstaden, France) was used as transfection reagent. Briefly, DNA:PEI complexes were formed by adding PEI to plasmid DNA diluted with supplemented HyCell™ media and vortexed for 10 seconds. PEI was added in 1:2 (w/w) DNA:PEI ratio and vortexed three times, for 3 seconds, incubated for 15 minutes at room temperature and then added to the cell culture. Polyplexes represented a 10% of the total culture volume to be transfected.

For the medium exchange (ME) condition, cell culture was centrifuged 48 hpt at 300 x g for 5 minutes, cell supernatant was harvested, and the cell pellet was gently resuspended with fresh media. All conditions were performed in triplicate.

For rAAV harvest, cell culture samples were taken at 48, 72, and 96 hpt. For cell supernatant (SN) collection, cell culture sample was collected by centrifuging the culture at 1000 x g for 10 minutes. For cell lysate collection, whole cell culture was treated with 1x cell lysis solution (500 mM Tris (Sigma-Aldrich, St. Louis, MO, USA), 1 % Triton X-100 V/V (Sigma-Aldrich, St. Louis, MO, USA), 20 mM $MgCl_2$ at pH 7.5) and incubated at 37ºC, 130 rpm, for 1 hour. Lysate mixture was centrifuged

at 4000 x g for 20 minutes and supernatant was collected as cell lysate (Lys) fraction. All fractions were stored at -80ºC until analysis was performed.

Infection for adenovirus vector (Ad-GFP) production

AdV5 E1AΔ-GFP (AdTLRGDK, Ad-GFP hereafter) was kindly provided by Theriva™ Biologics (Parets del Vallès, Spain)[364]. Ad-GFP stocks were diluted with PBS to infect cells at a multiplicity of infection (MOI) of 10 at a cell density of 2 x 10^6 cells/mL. Transfections were carried out 4 hpi (hours post infection) with either pMock or pE1_cnt at 1 µg/ml of total DNA. DNA:PEI complexes were formed as previously described. To perform medium exchange, cell culture was centrifuged at 24 hpt or 24 hpi as previously described. For Ad-GFP harvest, cell culture samples were taken at 24, 48 hpt/hpi. Cell lysate was collected as previously described.

Quantitative polymerase chain reaction (qPCR) for rAAV quantification

For the determination of physical vector genome titer of rAAV (vg), 5 µL of the cell lysate fraction or cell supernatant fraction from each sample was added to DNase I (0.066 mg/mL, 10104159001, Roche, Basel, Switzerland) prepared in DNase I buffer (0.16 mM $CaCl_2$, 6.6 mM $MgCl_2$, 17.3 mM Tris-HCl, 0.066 % Pluronic F68, pH 8.0). This mixture was incubated at 37 ºC for 16 hours. Then, 4 µL of 30 mM EDTA (Invitrogen, Carlsbad, CA, USA, in 10 mM Tris-HCl, pH 8.0) was added to each tube and incubated at 70ºC for 30 minutes to inactivate DNase I. After, 5 µL of Proteinase K (EO0491, ThermoFisher Scientific, San Jose, CA, USA) was added, incubated at 55ºC for 2 hours and inactivated at 95ºC for 15 minutes. The quantification of viral genomes was conducted using quantitative PCR (qPCR). This reaction was set up using iTaq Universal SYBR Green Supermix (BioRad, Hercules, CA, USA) and a set of primers specific to a region within the eGFP gene (**Table 1**). Thermocycling profile was based on manufacturer's recommendations in a CFX384 Touch Real-Time PCR Detection System (**Table 2**). For preparing the standard curve, a gBlock from IDT (Integrated DNA Technologies, Leuven, Belgium) containing the targeted eGFP region of known concentration was used (sequence shown in **Table 1**). A serial dilution of this standard, ranging from 10^8 to 10^2 copies/µL was prepared and amplified by qPCR. The concentration of AAV samples was the determined by interpolating in a standard curve. Additionally, an internal rAAV1-CMV-GFP reference was amplified alongside the samples in all the quantifications.

To quantify the accumulated rAAV titer in the supernatant (SN) (extracellular media), **Equation (1)** was used. To quantify the rAAV accumulated titer in the whole cell culture (cell lysate (Lys) fraction), **Equation (2)** was used. The total accumulated rAAV in the supernatant (Acc. rAAV in SN) was determined by adding the titer of the supernatant fraction at harvest (72 hpt) to the titer of the

harvested supernatant at 48 hpt. Similarly, the accumulated rAAV in the cell lysate fraction (Acc. rAAV in Lys) was determined by adding the titer of cell lysate fraction at harvest (72 hpt) to the titer of the harvested supernatant fraction at 48 hpt.

$$\text{Acc. rAAV in SN (vg)} = \text{rAAV titer in SN 72 hpt} \times V + \text{rAAV titer in SN 48 hpt} \times V \quad \textbf{(1)}$$

$$\text{Acc. rAAV in Lys (vg)} = \text{rAAV titer in Lys 72 hpt} \times V + \text{rAAV titer in SN 48 hpt} \times V \quad \textbf{(2)}$$

To determine the rAAV egress rate (referred to as Esn), the proportion of rAAV present in the supernatant fraction relative to the total rAAV was calculated using **Equation (3)**:

$$\text{Esn (\% SN)} = \frac{\text{Acc. rAAV in SN}}{\text{Acc. rAAV in Lys}} \times 100 \quad \textbf{(3)}$$

qPCR for Ad-GFP quantification

For the determination of physical vector genome titer of Ad-GFP, standard protocol as described for rAAVs was used with the following modifications. DNase I (10104159001, Roche, Basel, Switzerland) was added to lysate samples at 0.066 mg/mL and incubated at 37ºC for 16 hours in DNase buffer (5 mM $CaCl_2$, 5 mM $MgCl_2$, 50 mM Tris-HCl, pH 8.0). A purified Ad-GFP treated in parallel using the same protocol was used as internal reference and serially diluted to generate a standard curve. Ad-GFP samples concentration was determined interpolating in the standard curve.

Table 2. Conditions of cycling temperatures and time for qPCR and RT-qPCR**.**

Reaction	Cycles (temperature and time)
qPCR for rAAV quantification and RT-qPCR	95ºC, 3 min + 40 x [95ºC, 10 sec + 60°C, 30 sec] + 95ºC, 10 sec + 65 to 95°C, 0.5ºC/5 sec + 4°C hold

Reverse transcription quantitative PCR (RT-qPCR)

Total RNA was extracted from cell pellets kept at -80ºC using the Promega Maxwell RSC simplyRNA Tissue Kit (Cat No. AS1340) according to the manufacturer's protocol. Quantity of isolated RNA was assessed using Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, San Jose, CA, USA). cDNA was synthesized using 1 µg of total RNA per sample with the iScript cDNA Synthesis Kit (BioRad, Hercules, CA, USA).

Quantitative real time PCR (RT-qPCR) assays were performed with iTaq Universal SYBR Green Supermix (BioRad, Hercules, CA, USA) CFX384 Touch Real-Time PCR Detection System. Each

qPCR reaction mixture contained 2 µl of 5x diluted cDNA and a primer concentration of 500nM. The qPCR primers and program used for both target gene and housekeeping tested genes are provided in **Table 3**. At the end of the PCR, a melting curve was recorded from 65 to 95ºC. The qPCR assay was performed in triplicate for all samples. Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method[365]. Briefly, the target gene Ct values for each sample were subtracted the mean Ct value of housekeeping gene 18S of each sample to obtain ΔCt. Then, ΔΔCt was calculated by substracting ΔCt of the target gene to the ΔCt of reference control sample. Fold changes in relative gene expression were calculated by $\log_2(2^{-\Delta\Delta Ct})$. All conditions were performed in triplicate. Primer sequences are provided in **Table 3**.

Table 3. List of primers for RT-qPCR. Primers used to characterize housekeeping genes are also depicted.

RT-qPCR primers	Sequence (5'-3')
Fwd_E1a_total	TCCGGTCCTTCTAACACACC
Rev_E1A_total	GGCGTTTACAGCTCAAGTCC
Fwd_18S	TCAAGAACGAAAGTCGGAGG
Rev_18S[(a)]	GGACATCTAAGGGCATCACA
Fwd_UbC	GCCGCACTCTTTCTGACTACAAC
Rev_UbC	ACCTCCAGAGTGATGGTCTTGC
Fwd_EEF1A1	CCTGCCTCTCCAGGATGTCTAC
Rev_EEF1A1	GGAGCAAAGGTGACCACCATAC
Fwd_RPL22	TGATTGCACCCACCCTGTAG
Rev_RPL22	GGTTCCCAGCTTTTCCGTTC
Fwd_Rplp0	GAAACTCTGCATTCTCGCTTCC
Rev_Rplp0	GACTCGTTTGTACCCGTTGATG
Fwd_TBP	CGGCTGTTTAACTTCGCTTCC
Rev_TBP	TGGGTTATCTTCACACGCCAAG
Fwd_YWHAZ	AGGAGATTACTACCGTTACTTGGC
Rev_YWHAZ	AGCTTCTTGGTATGCTTGTTGTG
Fwd_USP25	CGGCGCAGAAGCACCAG
Rev_USP25	TCGCAGTAAGGAAAGCCACT
Fwd_β-actin	CACCATTGGCAATGAGCGGTTC
Rev_β-actin	AGGTCTTTGCGGATGTCCACGT

[(a)] From Fromigué *et al.*[366].

rAAV purification

Cell lysate was treated with Benzonase® (Merck Millipore, Kankakee, IL, USA, 25 U/mL, 1 mM Tris, 0.20 mM $MgCl_2$, pH 8) in 5 mL of volume for 1 hour at 37ºC, 130 rpm to reduce DNA contamination. After 30 minutes of incubation, $MgS0_4$ was added at 0.0375 M to decrease rAAV aggregation. Then, 1/10 of the volume of chloroform was added and the sample was vortexed for 5 minutes. After that, NaCl was added to each tube at a final concentration of 1M and vortexed for an additional 10 seconds. Samples were centrifuged 5 minutes at 3000 x g and aqueous phase was collected and transferred into 50 mL conical tube. To further reduce protein impurities, chloroform was added at 1:1 (v:v) to each tube, vortexed for 10 seconds and centrifuged 5 minutes at 3000 x g twice until aqueous phase was clear. Aqueous phase was collected, and tubes were left open in a biosafety cabinet to evaporate chloroform for 45 minutes. Sample was then transferred to Amicon Ultra-4 Centrifugal Filter (Merck Millipore, Kankakee, IL, USA) and tubes were centrifuged 5 minutes at 4000 x g. Then, elute was discarded, and 1 mL of PBS was added to the retentate for sample buffer exchange through centrifugation. This step was repeated twice to wash the concentrate. Finally, samples were sterile filtered with 0.22 µm and stored at -80ºC until analysis was performed. Samples were treated in triplicate and kept in ice during all the purification procedure. Centrifugations were performed at 4 ºC.

rAAV1 *in vitro* potency assay

HEK293 cells were seeded at 1 x 10^5 cells/well in a 96-well plate and incubated for 1 hour at 37ºC until cells were adhered to the plate. rAAV1 preparations were tested at various multiplicities of infection (MOIs) ranging from 100 to 400 yielding less than 50% GFP-positive cells to avoid inaccuracies derived by multiple viral particles infection of a single cell[352,353]. To ensure comparability, titers were normalized to 1.2 x 10^9 vg/mL in all samples before transduction. Cells were harvested 48 hours post-transduction (hptd) and the percentage of GFP-positive cells was assessed using Cytoflex (Beckman Coulter Inc., Brea, CA, USA) at Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria (UAB, Bellaterra, Spain). First, SSC-H vs FSC-A and FSC-A vs FSC-H density plots were generated to gate individual cell population. Then, GFP-positive population was assessed in Y585-PE-A vs GFP B525-FITC-A density plot. Gating was manually adjusted with HEK293 non-transduced cells as a control. Samples were concentrated 2-fold before analysis. 5000 events were analyzed per sample at a flow rate of 70 µL/min. Results were analyzed with the CytExpert v.2.3 software. The TU was calculated using the following expression (**Equation 4**):

$$\text{Transducing units (TU)/mL} = \frac{\text{GFP} \times \text{N} \times \text{FD} \times 1000}{\text{V}} \quad \textbf{(4)}$$

where GFP is the number of GFP-positive cells per well, N is the number of cells per well at the time of transduction (corresponding to 1 x 10^5 cells per well), FD is the fold dilution of vector sample used for transduction at different MOI's, V is the volume of diluted vector sample added into each well for transduction in microliters[352,353].

Western Blot of E1A protein

For Western Blot analysis, 600 µL of cell culture was centrifuged 300 x g for 5 minutes. Supernatant was discarded and cell pellet was resuspended in PBS and 1x cell lysate buffer for 1 hour at 37 ºC 130 rpm. Samples were centrifuged for 20 min at 4000 g at 4ºC and supernatant was then transferred into a clean tube on ice. Samples were treated with 1x Laemmli Sample Buffer (BioRad, Hercules, CA, USA), mixed and boiled for 20 minutes at 96ºC. Samples were ran on precast SDS-polyacrylamide (4–20%) gel electrophoresis at 200 V, 400 mA for 45 minutes. Precision Plus Protein Western C (#1610376, BioRad, Hercules, CA, USA) was used as molecular weight standard. Running buffer used was Tris/Glicine/SDS (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) (BioRad, Hercules, CA, USA). Gel was transferred onto a polyvinylidene difluoride (PVDF) membrane for 7 minutes using the system Trans-Blot Turbo Transfer System (BioRad, Hercules, CA, USA). Membrane with transferred proteins were incubated for 1 hour in blocking buffer (1x PBS, 0.1% (v/v) Tween-20 with 5% (w/v) non-fat milk) followed by 1 hour incubation with the corresponding primary antibody at room temperature with gentle shaking. Primary antibodies used were mouse anti-E1A monoclonal antibody (clone M73, Merck Millipore, Kankakee, IL, USA, 1:100) and mouse anti-β-actin antibody (MA5-15739, ThermoFisher Scientific, San Jose, CA, USA, 1:5000) was used as loading control. Membrane was washed three times for 5 minutes on orbital shaker with wash buffer (1x PBS, 0.1% (v/v) Tween-20). After primary incubation, a secondary incubation was performed using HRP-linked anti-mouse (A3562, Sigma-Aldrich, St. Louis, MO, USA) at 1:2500 in 2.5% (w/v) nonfat dry milk 1x PBS 0.1% (v/v) Tween-20 for 1 hour at room temperature. Proteins were visualized using Clarity Western ECL Substrate solution (#1705061, BioRad, Hercules, CA, USA) using ChemiDoc MP Imaging System (BioRad, Hercules, CA, USA).

Total capsid determination by ELISA

AAV capsid titers were determined by ELISA (RAAV Titration ELISA, cat# PRAAV1, PRAAV6, PRAAV8, PRAAV9, PROGEN Biotechnik GmbH, Heidelberg, Germany) for each different serotype tested, according to the manufacturer's instructions. Samples were serially diluted to fit the

standard curve of reference control provided by the kit. The standard curve was generated with the standard provided kit using a second-order quadratic polynomial curve fit and sample concentrations were automatically interpolated from the standard curve. Absorbances were acquired at 450 nm (SPECTROstar Nano, Ortenberg, Germany). Concentration values were averaged across the dilutions that were within the adjusted absorbance range recommended by the vendor and the standard deviation calculated. A minimum of two and to up to three dilutions were used for each sample[341].

Transmission electron microscopy (TEM)

rAAV1 preparations were analyzed after air-dried negative staining. Briefly, rAAV1 samples from NaCl-chloroform purification protocol were deposited onto carbon-coated copper or Holey carbon 200 mesh grids (Micro to Nano, Wateringweg, the Netherlands). Grids were glow discharged in a PELCO easiGlow glow discharge unit (PELCO, Fresno, CA, USA). Thereafter, 8 µL of sample were loaded onto the grid and incubated at room temperature for 1 min. Excess sample was carefully drained off the grid with the aid of filter paper. Samples were negatively stained with 8 µL of 2% w:v uranyl acetate by incubation at room temperature for 1 minute[367]. Excess stain was drained off as previously indicated and grids were dried at room temperature until analysis. TEM examinations were performed with a Jeol JEM-1400 (JEOL USA, Pleasanton, CA, USA) transmission electron microscope equipped with an ES1000W Erlangshen charge-coupled device camera (CCD) (Model No. 785; Gatan, Pleasanton, CA, USA) at Servei de Microscòpia (UAB, Bellaterra, Spain).

Total protein quantification

Host cell protein concentration was determined with the bicinchoninic acid assay using the Micro BCA protein assay kit (#23225, ThermoFisher Scientific, San Jose, CA, USA) according to the manufacturer's instructions. Briefly, serial PBS-buffer dilutions (between 2-and 256-fold) of reference standard and samples were dispensed into the wells. After adding 150 µL of the Micro BCA working reagent into each well, plates were incubated for 1 hour at 37°C. The bovine serum standard curve ranged from 1.5 to 200 µg/mL, and the absorbance was measured at 562 nm in the SPECTROstar reader (Ortenberg, Germany). Protein concentration of the samples was determined interpolating the reference standard curve.

Statistical analysis

All data were represented as mean ± standard deviation. Unless stated, one-way analysis (ANOVA) was used for all statistical analysis in this study. GraphPad was used to conduct statistical analysis. *p* values less than 0.05 were considered statistically significant.

RESULTS

Endogenous Adenoviral E1 expression in HEK293 cells fluctuates across different rAAV production protocols

In **Chapter one**, by means of a proteomic analysis, individual proteins were identified and quantified showing differential changes at different time points (48, 72, 96 hpt) across three conditions including a growth condition and two rAAV production conditions: standard triple transfection (TT) and extended gene expression (EGE). The EGE was performed by a medium exchange (ME) and an additional re-transfection (RT) at 48 hpt.

As described, different strategies of rAAV9 production impacted on the adenoviral E1 protein expression in HEK293 cells. Here, individual changes in E1A and E1B protein levels were characterized. Remarkably, the E1A 289R protein was downregulated in TT method in comparison to growth condition, showing a marked decrease at 96 hpt. In contrast, at the same time point, E1A protein levels in the EGE protocol were similar to those observed in the growth control condition (**Figure 1A**). This suggests that ME, RT, or a combination of both, may play a role in preserving E1A expression to levels observed in normally proliferating cells.

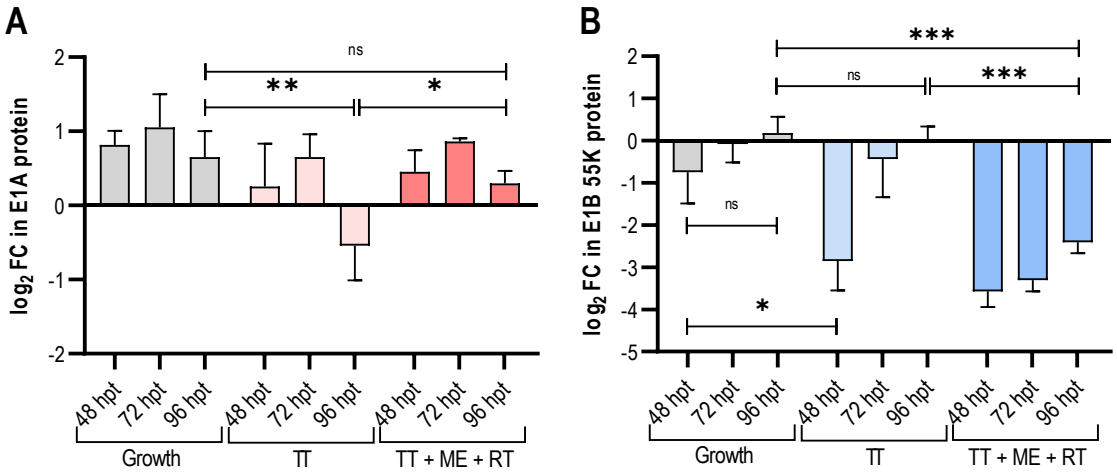


Figure 1. Analysis of adenoviral E1 gene expression changes in HEK293. Proteomic data obtained in Chapter one represented with bar plots showing average log_2FC values of each protein in comparison to peptide pool corresponding to growth condition at 48hpt. **A.** E1A protein (UniProt: P03254). **B.** E1B 55K protein (Uniprot: P03255). ME and RT were performed at 48 hpt, after sample extraction at this time point. Combination of TT+ME+RT results in the EGE method. Biological triplicates were measured in triplicates. Data was analyzed by Ordinary ONE-way ANOVA *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. ME, media exchange. RT, re-transfection. TT, triple transfection.

Further, E1B 55K showed a decrease due to rAAV9 production that was restored over time in TT. In contrast, the EGE protocol hindered E1B 55K expression recovery and maintained a lower E1B 55K expression compared to growth condition at the same time point (**Figure 1B**).

The role of adenoviral E1A protein has been extensively identified as a key factor required for rAAV production[101]. Thus, the downregulation observed in E1A is hereby studied as a putative bottleneck in rAAV production using HEK293 cells, hypothesizing that the E1A expression preservation during the EGE method may be associated with enhanced viral particle yield.

Media exchange enhances rAAV9 and rAAV1 titer and viral egress

Having identified that the outperformance of the EGE protocol compared to TT production may be associated with E1A expression preservation, this study first aimed at describing the specific effects involved in the EGE method on rAAV yield. To this end, the experimental design included: a ME after transfection (ME) performed at 48 hpt, a re-transfection (RT) at 48 hpt, and a complete EGE involving both ME and RT. The standard TT without ME or RT was included as control. First, the study focused on rAAV9 since this serotype was used to perform the proteomic analysis. Then, translatability to other serotypes (rAAV1) was evaluated (**Figure 2A-F**).

Of note, rAAV titer was measured based on the accumulated titer, which includes the rAAV collected in the cell culture supernatant at 48 hpt by the ME step. This provides an estimation of the accumulated titer as if a continuous harvest in a perfusion setting was performed. For the quantification of the accumulated titer at cell supernatant fraction, **Equation 1** was applied. Similarly, the accumulated titer in the cell lysate fraction, which represents the total rAAV present in the cell culture was assessed using **Equation 2**. Additionally, by quantifying rAAV in both cellular fractions, the efficiency of rAAV egress under each set of conditions was determined (**Equation 3**).

Upon analysis of the rAAV9 titer in the cell lysate fraction, no significant differences in rAAV9 titer between RT and TT were observed. However, ME led to a slight increase in yield, comparable to that achieved by the EGE protocol (**Figure 2B**). Further, examination of the supernatant fraction showed no significant increase with the RT compared to the TT protocol (**Figure 2A**). In contrast, the application of ME alone, resulted in a significant increase in extracellular rAAV9, similar to the enhancement provided by the EGE protocol (**Figure 2A**). This suggests that the predominant enhancement effect conferred by the EGE protocol can be attributed to the ME step.

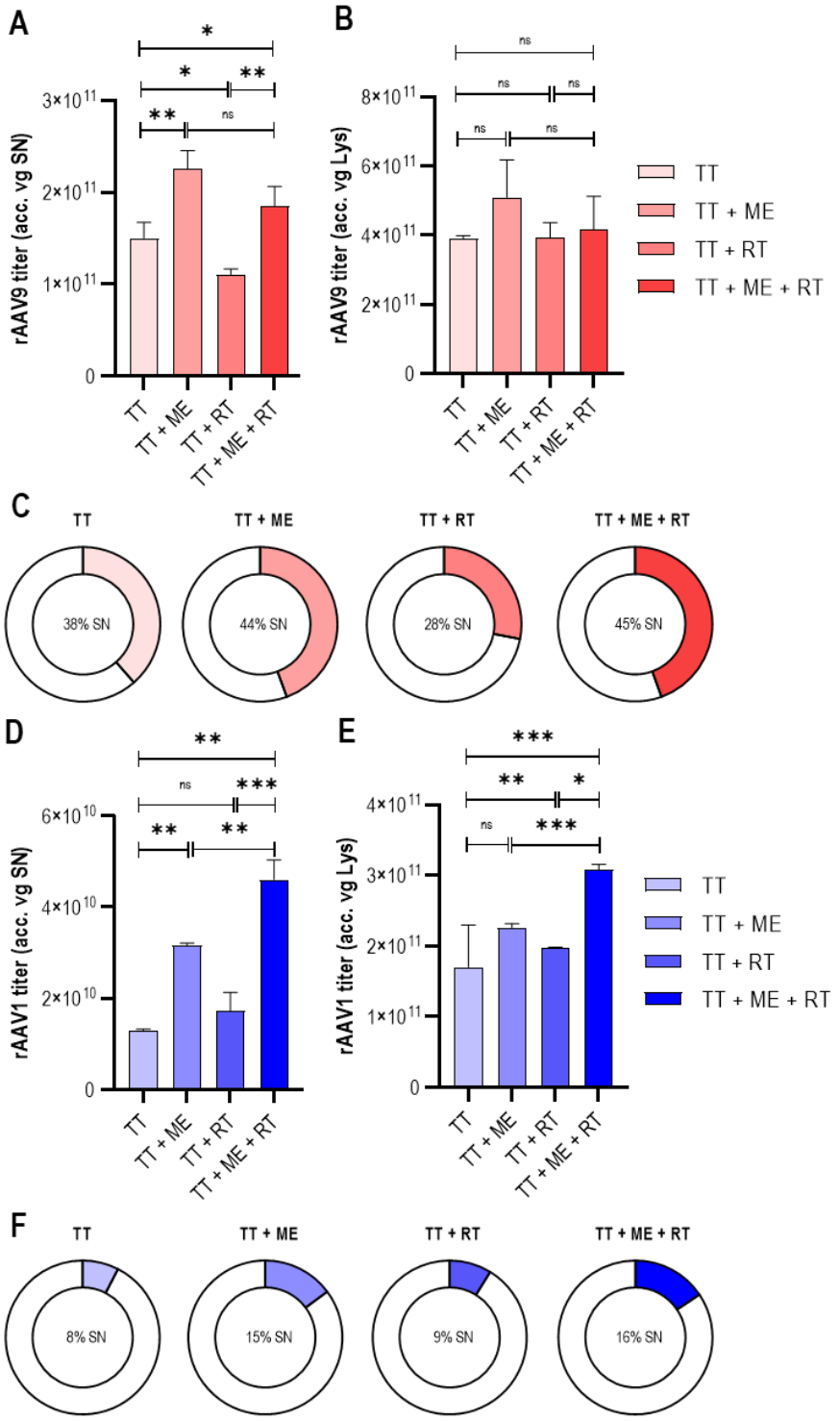


Figure 2. Impact of media exchange, re-transfection and extended gene expression on rAAV9 and rAAV1 production. A. Accumulated rAAV9 harvested in cell supernatant**. B.** Accumulated rAAV9 harvested in cell lysate. **C.** Percentage of rAAV9 recovered in cell supernatant from total production. **D.** Accumulated rAAV1 harvested in cell supernatant**. E.** Accumulated rAAV1 harvested in cell lysate. **F.** Percentage of rAAV1 recovered in cell supernatant from total production. ME, RT and EGE protocols were performed at 48hpt. Biological triplicates were measured in triplicates. Data was analyzed by ANOVA analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. EGE, extended gene expression. Lys, lysate. ME, medium exchange. RT, re-transfection. SN, supernatant. TT, triple transfection. Vg; vector genomes.

This is further supported by the analysis of the distribution of rAAV9 in the supernatant relative to the total production. While the TT control condition yielded 38% of the viruses in the supernatant, only 28% were recovered in the RT condition. Significantly, the ME resulted in 44% of rAAV9 exiting the cells, comparable to the 45% observed with the complete EGE protocol, underscoring ME's significant role in promoting viral egress (**Figure 2C**).

To assess the applicability of these findings across different AAV serotypes, rAAV1 production under the same conditions was analyzed. In the cell lysate fraction, RT did not increase rAAV1 titer. Furthermore, while ME alone did not significantly boost rAAV1 production, the combination of ME and RT within the EGE protocol markedly increased viral yield (**Figure 2E**). In the cell supernatant fraction, RT alone did not increase viral yield; however, ME significantly elevated the extracellular rAAV1 titer, further enhanced by the complete EGE protocol (**Figure 2D**).

Finally, the percentage of viral egress for rAAV1 under each condition was evaluated. Here, ME alone increased rAAV1 recovery in the supernatant fraction, improving egress from 8% in the TT control to 15-16% with and without re-transfection (ME and EGE), respectively (**Figure 2F**).

Overall, the implementation of ME and EGE strategies not only resulted in increased accumulated rAAV yield in the cell supernatant for both rAAV9 and rAAV1 but also enhanced the rate of viral egress, demonstrating the potential of these protocols to optimize rAAV production processes. Additionally, serotype-dependent differences were observed in the efficacy of the EGE method. Specifically, the EGE remarkably increased rAAV1 specific production. In contrast, this method did not significantly enhance rAAV9 production.

Media exchange exerts a preservation effect on endogenous E1A gene expression in HEK293 cells during rAAV production

Having determined that ME is the key component in the EGE protocol that increases rAAV egress, we next sought to elucidate the specific effects of each component within the EGE protocol (ME and RT) that contribute to preserving E1A protein expression at levels comparable to the cell growth condition at 96 hpt. To this end, E1A expression during rAAV9 and rAAV1 production was analyzed at 96 hpt by real-time qPCR (RT-qPCR) and Western Blot (**Figure 4**).

To assess significative changes in mRNA expression levels during rAAV production, a stable reference gene (housekeeping gene) was first identified to provide a reliable baseline to evaluate changes in gene expression. This proved to be challenging as an overall disruption of host transcriptional processes during rAAV production was observed[1]. To address this, the expression variation of nine different housekeeping genes (provided in **Table 3**) between cell growth and TT condition at 96 hpt was measured and the ΔCt values of TT compared to cell growth were analyzed[368] (**Figure 3**). Among the tested reference genes, the expression levels varied remarkably upon TT production for all genes except for UbC, TBP, and ribosomal protein 18S. Among the those, the ribosomal 18S gene was selected as a reliable housekeeping gene for rAAV gene expression in subsequent analyses.

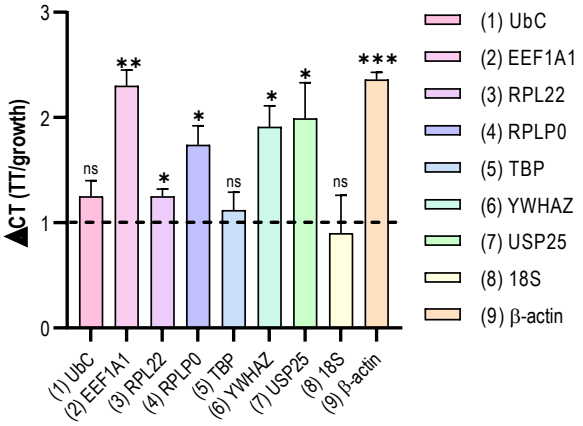


Figure 3. Evaluation of housekeeping gene stability during rAAV production. Analysis of changes in mRNA expression levels of various housekeeping genes during rAAV production compared to cell growth condition at 96 hpt. The ΔCt values from TT compared to cell growth were compared. Biological triplicates were measured in triplicates. Data was analyzed by t-test in comparison to 1 set as control. *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD.

Initially, validating previous findings, a significant reduction in E1A expression was observed during the TT protocol for rAAV9 production, with both mRNA and protein levels decreasing markedly compared to the cell growth control (**Figure 4A, B**). This downregulation was also observed in rAAV1 production using the standard TT (**Figure 4C, D**). Additionally, upon the RT of the cell culture, a decrease in E1A expression was also observed across the different serotypes, both at mRNA level (**Figure 4A, C**), and protein level (**Figure 4B, D**). In contrast, following ME, E1A expression levels remained similar to those of the cell growth control, consistent for both rAAV9 and rAAV1 production (**Figure 4C, D**). When ME was integrated with RT as in the EGE protocol,

[1] In **Chapter one**, significant decrease on HEK293 transcription processes was observed over rAAV production in comparison to cell growth control condition, favoring viral transcription (see **Chapter one, Figure 3A and 3B**).

E1A expression was similarly preserved to levels comparable to the cell growth condition. These results suggest that the ME step within the EGE protocol plays a critical role in avoiding the E1A expression decrease during the production of both rAAV1 and rAAV9.

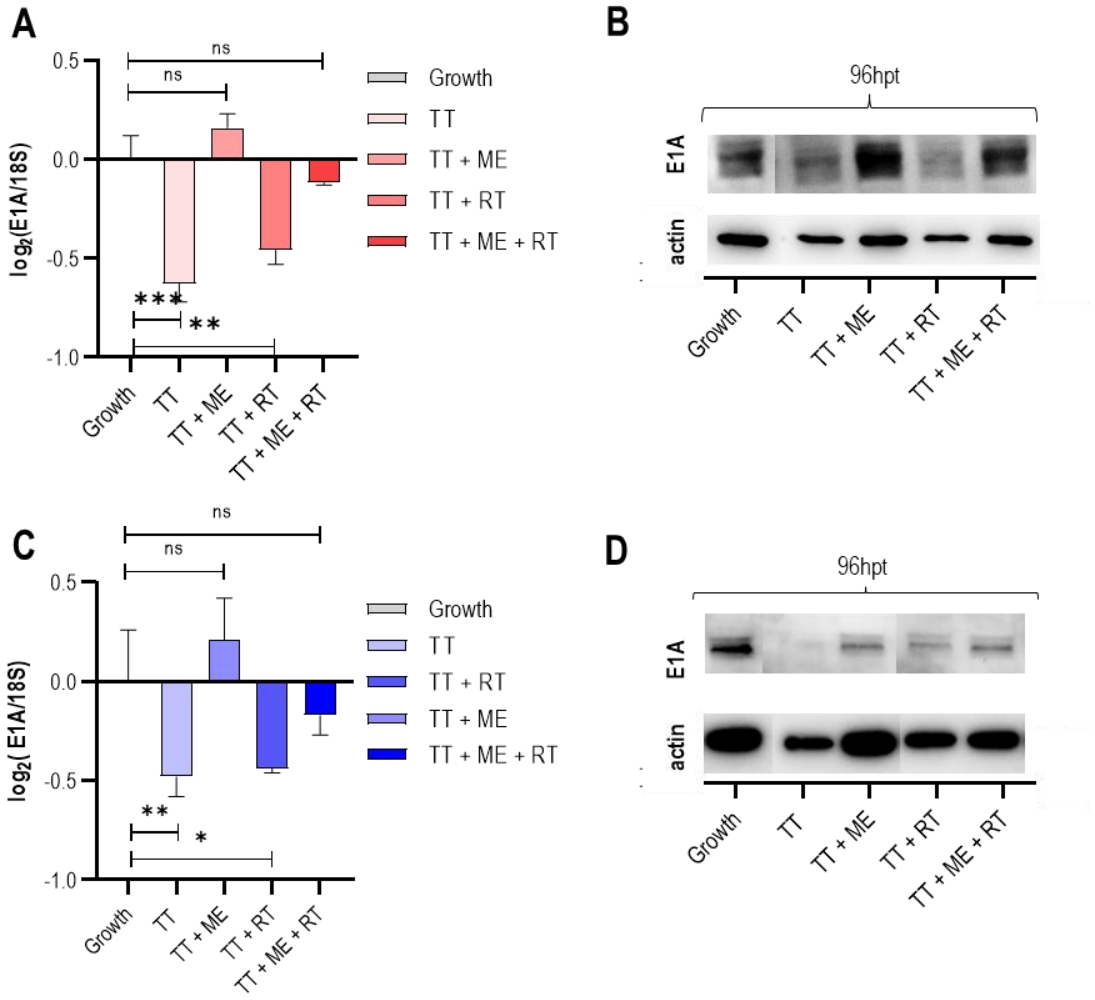


Figure 4. Characterization of E1A expression alterations in rAAV9 and rAAV1 production. This figure illustrates the effects of medium exchange (ME), re-transfection (RT), and the combined protocol (TT + ME + RT, referred to as EGE) on E1A protein and mRNA levels in rAAV1 and rAAV9 production, measured at 96 hours post-transfection (hpt). **A** and **C.** Total RNA transcript levels of E1A for rAAV9 and rAAV1, respectively, normalized to 18S RNA as the housekeeping gene. **B** and **D.** Western Blot images of E1A and actin protein expression 96hpt for rAAV9 and rAAV1, respectively. ME and RT were performed at 48hpt. Data was analyzed by ANOVA analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. TT, triple transfection. ME, medium exchange. RT, re-transfection.

Adenoviral E4 or VA RNAs expression triggers cell death and E1A expression decrease in HEK293 cells

Considering the previous findings, the element involved in E1A decrease during the TT protocol was investigated in HEK293 cells. Therefore, standard triple transfections were performed with one of the tested plasmids being excluded in each condition to assess its contribution to transfected cell cultures. Transfections were performed following the standard protocol and a transfection with pMock plasmid was included as control, as well as growth condition. All conditions were transfected using the same copy numbers of all plasmids and normalizing total DNA with stuffer DNA plasmid (pMock) to 1 µg/ml.

The effect on cell viability of the transfection process itself with an empty plasmid (pMock) or standard triple transfection plasmids was first evaluated. The cytotoxic effect of PEI in cell transfection is a well-described event, and in our results, transfection negatively impacted VCD compared to proliferating cells (**Figure 5A**)[207,316]. For rAAV1 production by TT, VCD was even more significantly decreased. Notably, when the pHelper plasmid was excluded from the transient transfection process, VCD was similar as in the transfection control condition with pMock, and higher than in TT method (**Figure 5A**). However, when pRepCap1 or pGFP_ITR plasmids were excluded from the TT, VCD was as negatively affected as in the TT condition (**Figure 5A**). This suggests that AdV helper genes contained in pHelper are the predominant key factors negatively influencing cell density at harvest in the triple transfection plasmids in the rAAV1 production process.

Regarding E1A expression dynamics, E1A mRNA and protein expression levels were not altered by pMock transfection, showing similar levels as growth condition (**Figure 5B, C**). This indicates that variations on E1A expression are unrelated to the cytotoxicity of transient transfection process itself. In addition, the decrease in E1A mRNA and protein expression observed with TT compared to the growth condition was also observed when packaging and GOI plasmids were excluded from TT transfection (**Figure 5B, C**). Interestingly, when pHelper plasmid was excluded, E1A expression levels were similar to pMock transfection and cell growth conditions (**Figure 5B, C**). These findings suggest that the pHelper plasmid is responsible for both the VCD reduction and the decrease in E1A expression.

Next, the specific AdV helper genes expressed within pHelper plasmid were evaluated regarding its toxicity and E1A expression variation potential. To this end, pHelper plasmid was split in two different plasmids: the pVA_RNA_E4 plasmid (pE4_VA hereafter), which contains the full E4 ORF and VA RNA I and II; and pE2, containing E2 ORF. Triple transfection was used as control as well

as quadruple transfection, using both pE4_VA and pE2 to supply pHelper functions. Remarkably, upon the exclusion of pE4_VA but not with pE2 exclusion, VCD was maintained (**Figure 5D**).

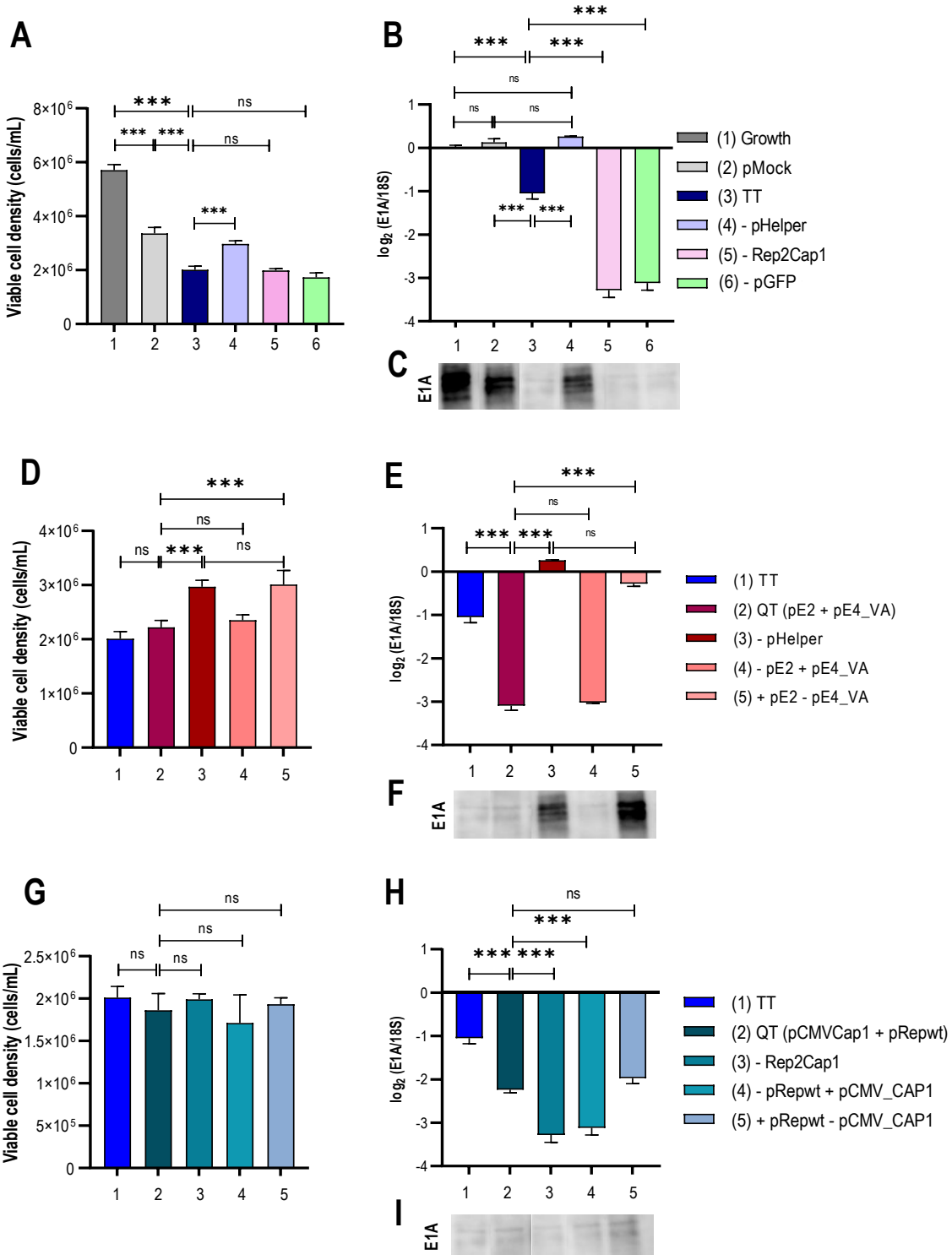


Figure 5. Analysis of elements involved in triple transfection for rAAV1 production on viable cell density (VCD) and E1A expression. A. D. G. Viable cell density upon different combinations of plasmid transfections as described in the text. **B. E. F.** E1A mRNA expression levels analyzed by RT-qPCR, using growth condition as control. **C. F. I.** Western Blot images of E1A protein. The amount of protein loaded was normalized to 200 µg/mL per lane. Data was collected at 96hpt and was analyzed by ANOVA analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. QT, quadruple transfection. TT, triple transfection.

Coherently, the apoptotic nature of E4orf4 has been already reported[128]. Interestingly, the exclusion of the pE4_VA plasmid in the quadruple transfection process resulted in E1A expression levels that were comparable to those observed when the pHelper plasmid was omitted. Conversely, omitting the pE2 plasmid led to a significant decrease in E1A expression (**Figure 5E, F**). These findings suggest that E4 ORF or VA RNAs are the responsible elements within pHelper plasmid for the reduction in both VCD and E1A expression during rAAV1 production.

In further experiments, the Rep and Cap ORFs of the pRep2Cap1 plasmid were separated to evaluate their individual effects. As expected, no differences in VCD were observed when pRep_wt or pCMV_CAP1 elements were tested compared to its correspondent quadruple transfection condition (**Figure 5G**). Finally, upon the exclusion of the *rep* gene, either in pRepCap1 or in pRep_wt plasmids (**Figure 5H**, condition 3 and 4), a more pronounced decrease in E1A expression was noted compared to quadruple transfection, without affecting VCD or protein levels as observed in the Western Blot analysis (**Figures 5H, I**).

The *E1A* gene expression preservation using medium exchange or its supplementation by metabolic engineering enhances rAAV production

Given the enhancement in rAAV production associated with the ME protocol, potentially partially explained by to the preservation effect on E1A expression, the effects of E1 overexpression through a metabolic engineering approach were evaluated. For this purpose, a pE1_cnt plasmid was constructed encoding the entire E1 region of wt AdV5 with its endogenous promoters (**Figure 6A**). Therefore, either pMock or pE1_cnt plasmid was co-transfected with the TT plasmids, maintaining the same copy number for each plasmid. The ME condition was kept as E1A expression preservation control during rAAV1 production.

Remarkably, the exogenous supplementation of E1 resulted in a two-fold improvement in accumulated extracellular titers and production in the cell lysate compared to the control condition (TT + pMock), reaching higher titers than those observed with the ME protocol (**Figure 6B, C**).

Additionally, an increase in E1A protein expression was observed with the supplementation of E1, slightly exceeding levels of the TT control condition (**Figure 6D**). Further examination of E1A total transcript levels showed that the pE1_cnt plasmid not only restored the downregulated E1A levels caused by rAAV1 production but also resulted in an overexpression of E1A (**Figure 6E**). These findings suggest that metabolic engineering achieved through exogenous addition of E1A enhances rAAV1 production efficiency, at higher rates than the ME protocol.

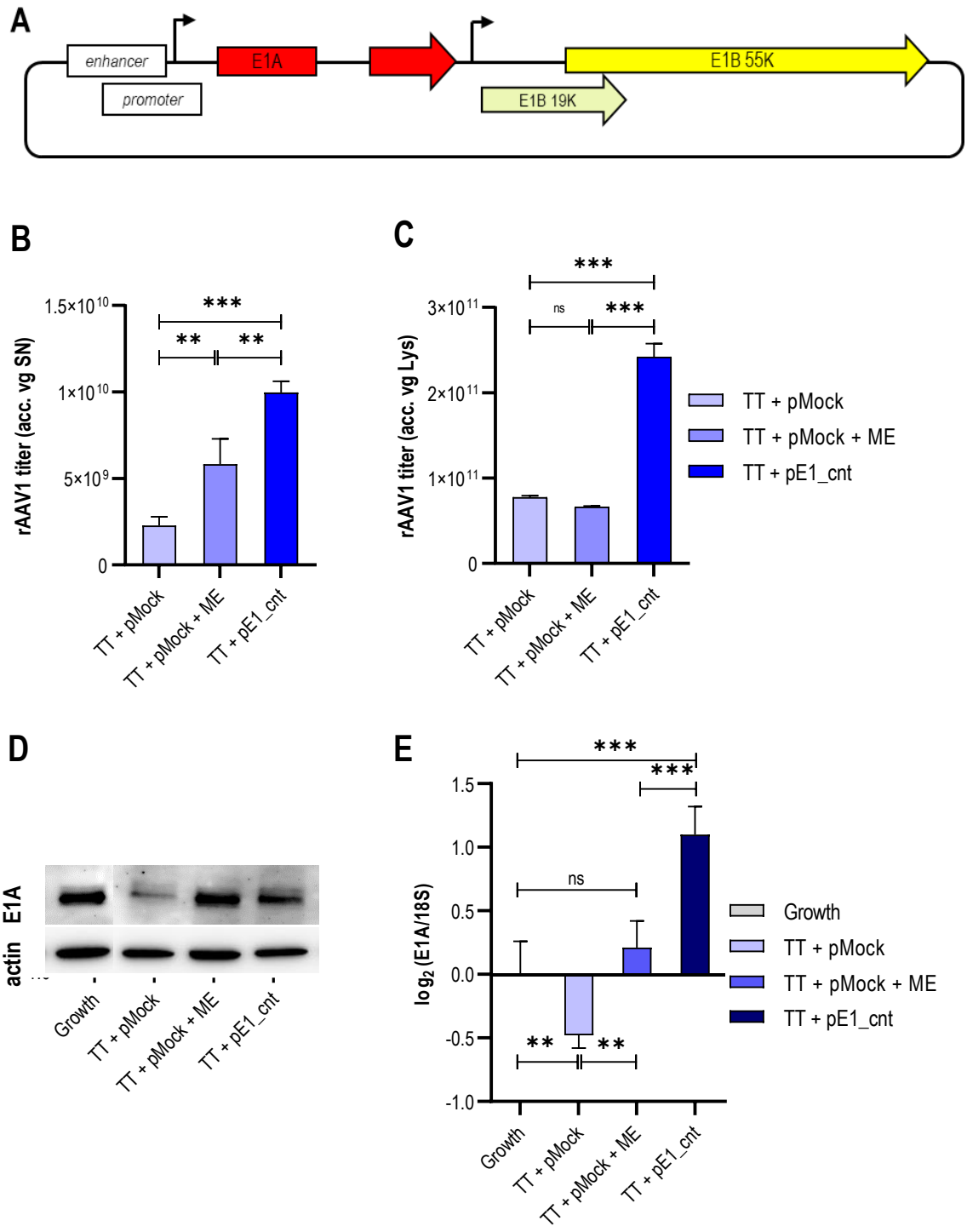


Figure 6. Characterization of E1A protein restoration using medium exchange or metabolic engineering approaches on rAAV1 production and E1A expression levels. A. Schematic representation of pE1_cnt plasmid showing coding region for E1A and E1B open reading frames. E1A enhancer and promoter are indicated. E1B 19K coding region partially overlaps with E1B 55K. **B.** rAAV1 production evaluated in cell supernatant. **C.** rAAV1 production evaluated in cell lysate. **D.** Western Blot showing E1A and actin protein expression at 96hpt. **E.** Total RNA transcript expression of E1A at 96hpt. ME was performed at 48hpt. Data was analyzed by ANOVA analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. Lys, lysate. ME, medium exchange. SN, supernatant.

E1A preservation strategies increase rAAV titer across different serotypes

After identifying that preserving or overexpressing E1A protein levels through either metabolic engineering or ME, respectively, during TT was beneficial for rAAV1 production, the transferability of these strategies was evaluated to other AAV serotypes, including rAAV6, rAAV8 and rAAV9. The accumulated titer using specific production in cell supernatant and cell lysate fraction at 72 hpt was quantified using **Equation 1** and **2** respectively.

Notably, the ME resulted in a remarkable enhancement in the cell supernatant titer for rAAV8 with a 141.2% (±31.3) increase, and slightly in rAAV6 and rAAV9 (**Figure 7A, C, E**). In the cell lysate fraction, rAAV6 experienced the highest increase (160.9 ± 61.0), followed by rAAV8 and a moderate effect in rAAV9 (**Figure 7B, D, F**).

E1 supplementation by metabolic engineering outperformed ME in rAAV1, rAAV6 and rAAV9, resulting in higher yields in both supernatant and cell lysate fractions (**Figure 7A-F**). Although metabolic engineering increased rAAV8 extracellular titer, it did not outperform ME in this specific serotype. These results demonstrate that the E1A increase through pE1_cnt supplementation is generally a more effective strategy compared to ME for enhancing rAAV production, although serotype-dependent yields in either supernatant or cell lysate fraction are observed. **Table 4** summarizes the percentage of increase observed using both strategies in the analyzed rAAV serotypes.

Table 4. Enhancement of rAAV titer in cell supernatant and cell lysate fraction across different serotypes using E1A preservation strategies. Percentage increase in accumulated rAAV titer at harvest (72 hpt) in both supernatant (SN) and lysate fractions for rAAV1, rAAV6, rAAV8 and rAAV9 using different strategies: media exchange (ME) performed 48 hpt and metabolic engineering by co-transfecting the TT with pE1_cnt plasmid. Control condition was triple transfection (TT) with pMock without media exchange.

rAAV serotype	Medium exchange (ME)		Metabolic engineering (pE1_cnt)	
	Increase in SN	Increase in cell lysate	Increase in SN	Increase in cell lysate
rAAV1	155.7 ± 85.1	-14.5 ± 2.5	337.6 ± 95.8	211.5 ± 10.1
rAAV6	57.6 ± 20.2	160.9 ± 61.0	130.1 ± 25.2	179.1 ± 43.6
rAAV8	141.2 ± 31.3	120.5 ± 18.5	39.1 ± 20.2	100.3 ± 17.0
rAAV9	50.7 ± 21.7	31.0 ± 28.0	110.0 ± 37.1	108.5 ± 45.1

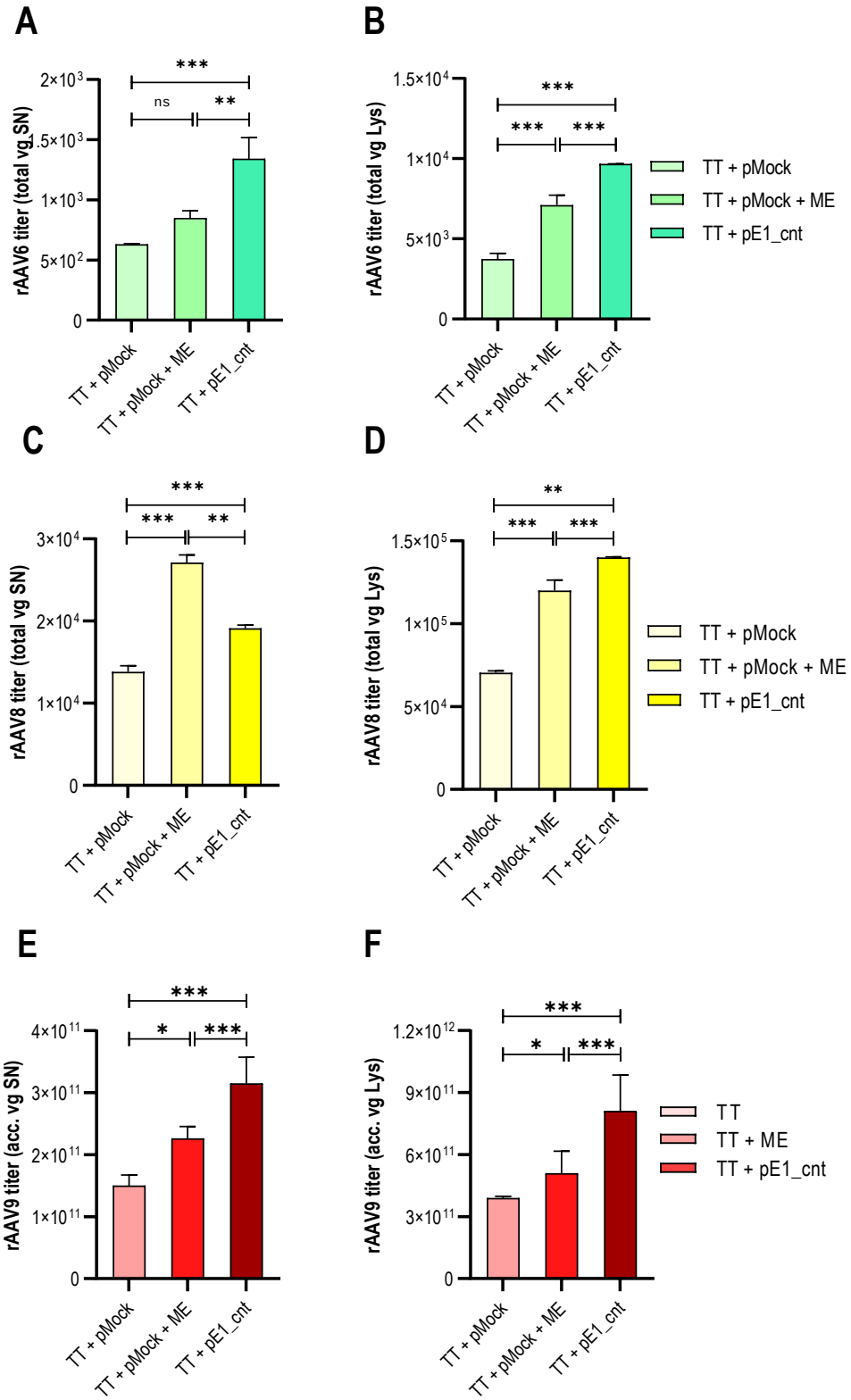


Figure 7. Characterization of accumulated rAAV titer across different serotypes with ME or metabolic engineering strategies. Accumulated titers of rAAV (total vector genomes, vg) in the cell supernatant and cell lysate fractions at 72 hours post-transfection (hpt) for various serotypes. **A, C**, and **E.** rAAV titers harvested from the cell supernatant for rAAV6, rAAV8, and rAAV9, respectively. **B, D**, and **F**. rAAV harvested from the cell lysate for rAAV6, rAAV8, and rAAV9, respectively. Titer calculations were performed using Equations 1 and 2 for the supernatant (SN) and lysate (Lys), respectively. Biological triplicates were measured in triplicates. Data was analyzed by ANOVA analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. Lys, lysate. ME, medium exchange. SN, supernatant. TT, triple transfection. Vg; vector genomes.

Potency is maintained over enhancement of rAAV1 titer by metabolic engineering

Although metabolic supplementation using pE1_cnt plasmid demonstrated an enhancement in rAAV production, biological activity is a critical quality attribute of rAAV preparations that needs to be evaluated as it can be affected depending on the production method[369,370]. This parameter measures the functional activity of the product resulting from vector transgene expression upon successful infection of target cells[352].

To assess the potency of the produced rAAVs using E1 supplementation strategy, the rAAV1-GFP vectors were produced and purified. Vectors produced by E1 supplementation by metabolic engineering approach were compared to the transient transfection production method using pMock. Importantly, the purity level may influence the rAAV transduction capacity[371]. Therefore, purification of rAAV1 was performed using an in-process compatible purification protocol adapted from Negreira *et al*., but avoiding PEG-precipitation step[372]. This modification enabled low-volume purification while maintaining high-throughput capacity. It also enabled the purification of biological triplicates in parallel, facilitating efficient comparison of the potency of the rAAV1 preparations by each production protocol.

To test the biological activity of the rAAV1-GFP vectors, a common biological assay was performed, consisting in measuring transduction capacity by quantifying transgene expression. HEK293 cells were used due to their high permissiveness to rAAV vectors. GFP expression was analyzed with fluorescence-activated cell sorting (FACS) 48 hours post transduction (hptd).

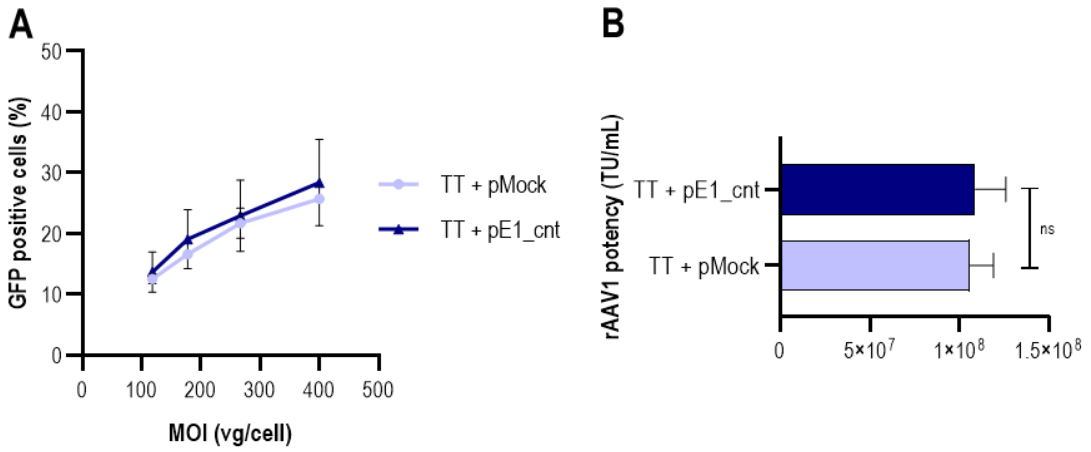


Figure 8. Titration of rAAV1-GFP Transducing Units by FACS analysis of GFP expression. A. HEK293 cells were transduced with rAAV1-GFP vector preparations from TT + pMock or supplementation by TT + pE1_cnt at the indicated MOIs (in vg/cell) and harvested 48 hptd. GFP positive cells (%) were analyzed by FACS. **B**. Transducing Units/mL (TU/mL) were calculated from different production methods. Bars represent the mean of TU/mL (calculated with Equation 4) (±SD) from 3 independent wells and three different infected MOIs. Results were analyzed using One-Way ANOVA. ns, $p > 0.05$; *$p < 0.05$. Data are presented as mean ± SD. MOI, multiplicity of infection. TT, triple transfection. Vg, vector genomes.

The percentage of GFP positive cells of the rAAVs derived from the two compared production methods (TT + pE1_cnt compared to TT + pMock) did not show significant differences in the range of the tested MOIs (**Figure 8A**).

The potency of the two rAAV production methods was compared based on transducing units per milliliter (TU/mL), calculated using **Equation 4**. Importantly, an equal transduction capacity reaching 1 x 10^8 TU/mL was observed comparing rAAV1 preparations produced with or without E1 supplementation via metabolic engineering (**Figure 8B**).

Combining strategies to increase E1A expression and rAAV production: the synergy of metabolic engineering and medium exchange

Two distinct approaches have been shown to enhance rAAV production: (1) performing a ME, which preserved E1A expression and (2) the addition of pE1_cnt plasmid to the TT production, thus providing additional levels of E1A from early stages of rAAV production. This study aimed to explore whether combining the pE1_cnt plasmid with media exchange could yield a synergistic effect on rAAV1 production.

Followed the method outlined in **Figure 9A,** all tested conditions initially underwent standard TT, followed by a ME at 48 hpt. Subsequently, RT was performed using the same copy numbers of all plasmids, organized in two different approaches: in one condition, RT was carried out only with pE1_cnt plasmid; in the other, RT was conducted using all TT plasmids. This combination of ME and RT is referred to as standard EGE and has been previously validated to enhance both supernatant and total cell-specific production (**Figure 2A and B**). A control condition RT with pMock was also included.

In the cell supernatant, no significant differences were observed in the three compared re-transfection conditions in rAAV1 production (**Figure 9B**). This finding indicates that the advantages achieved through ME does not further benefit from the RT process in terms of extracellular rAAV1 yield (**Figure 9B**).

Differently, rAAV1 production in the crude lysate improved by RT with pMock after ME, resulting in a slightly higher titer in comparison to ME. Additionally, RT with both TT and pE1_cnt yielded a similar increase in rAAV1 titer when compared to RT with pMock (**Figure 9C**). This suggests that the strategy of replacing spent media, followed by RT with either the standard TT or E1A supplementation via metabolic engineering achieves comparable enhancement of rAAV titer.

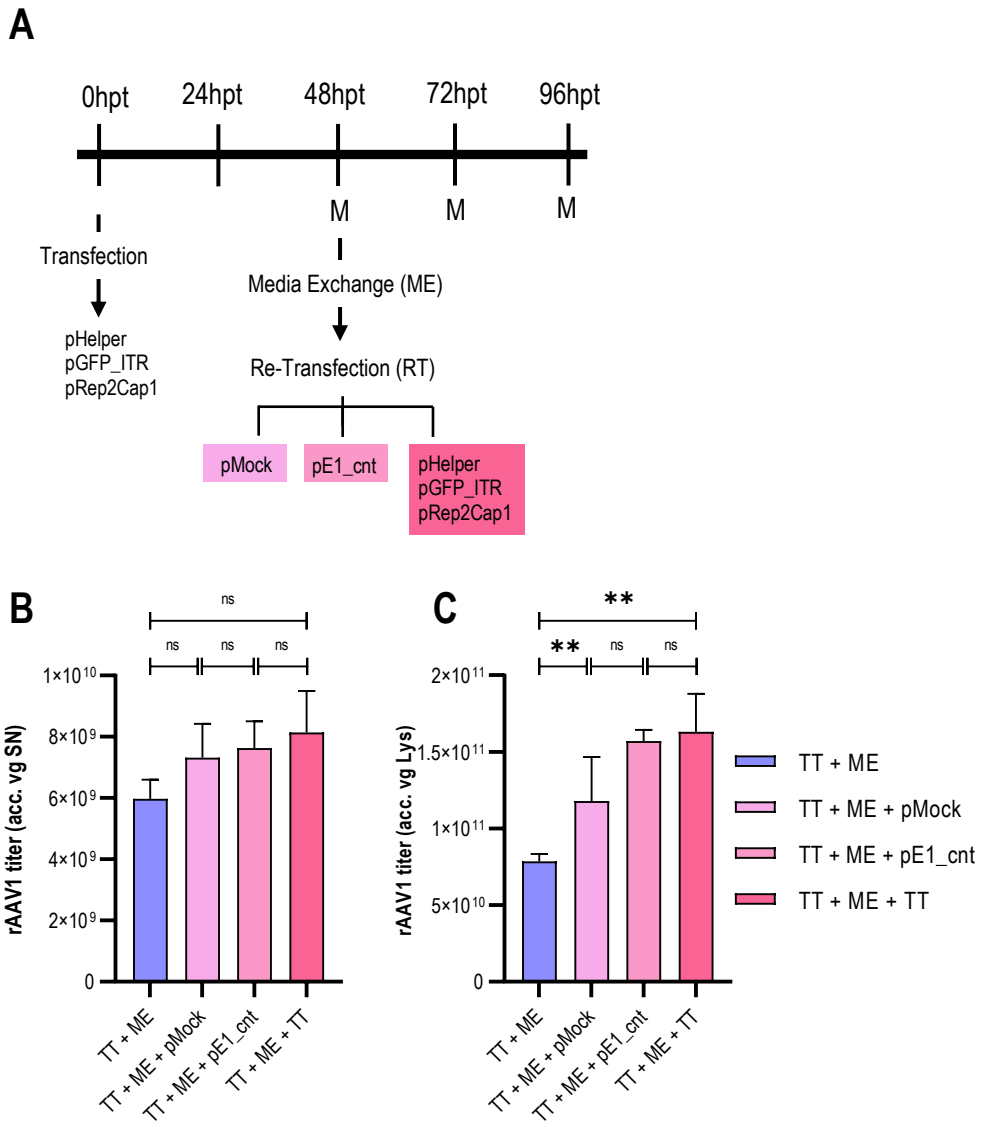


Figure 9. Combination of metabolic engineering by pE1_cnt plasmid and Medium Exchange on rAAV1 titer. A. Diagram of the different tested conditions. The combination of TT plasmids includes pHelper, pGFP_ITR and pRep2Cap1. After media exchange conducted 48hpt, re-transfection of pMock, pE1_cnt or TT was compared. Characterization of rAAV accumulated titer in cell supernatant and cell lysate at 72hpt. **B.** Accumulated rAAV harvested from cell supernatant. **C.** Accumulated rAAV harvested from cell lysate. Biological triplicates were measured in triplicates. ME was performed at 48hpt. Data was analyzed by ANOVA analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. EGE, extended gene expression. TT, triple transfection. ME, medium exchange. Vg; vector genomes.

E1A preservation strategies increase Ad-GFP production

As a *Dependovirus*, rAAV relies on the presence of AdV *E1A* gene to begin its replication. E1A is a key regulatory protein and plays a crucial role as a transcriptional activator, initiating viral life cycle. Consequently, similar to rAAV production using HE293 cells, E1A expression may be a limiting factor in adenoviral vector production.

To explore this, Ad-GFP vectors were produced by infection and ME was performed 24 hours post infection (hpi). To investigate the impact of E1A overexpression, transfection of either pE1_cnt or pMock plasmid was conducted at 4 hpi to avoid any potential interference with the Ad-GFP infection process[351].

First, E1A mRNA expression levels upon Ad-GFP vector production were evaluated in comparison to growth control condition. E1A expression levels decreased upon Ad-GFP vector production. However, upon ME, E1A expression levels were comparable to those of growth (**Figure 10A**). The E1A expression decrease was also observed for Ad-GFP infection followed by pMock transfection; similarly, following ME, E1A levels were again comparable to those seen in the cell growth (**Figure 10A**). With E1A overexpression by pE1_cnt plasmid, a remarkable increase in E1A expression was observed (**Figure 10A**).

The influence of E1A expression on Ad-GFP vector production yield was investigated. Significantly, an eight-fold increase in Ad-GFP vector titer in the cell lysate was observed upon ME post infection, indicating a substantial improvement in viral replication (**Figure 10B**). Additionally, supplementation with pE1_cnt plasmid resulted in a two-fold enhancement in viral vector production compared to the pMock control condition (**Figure 10A**).

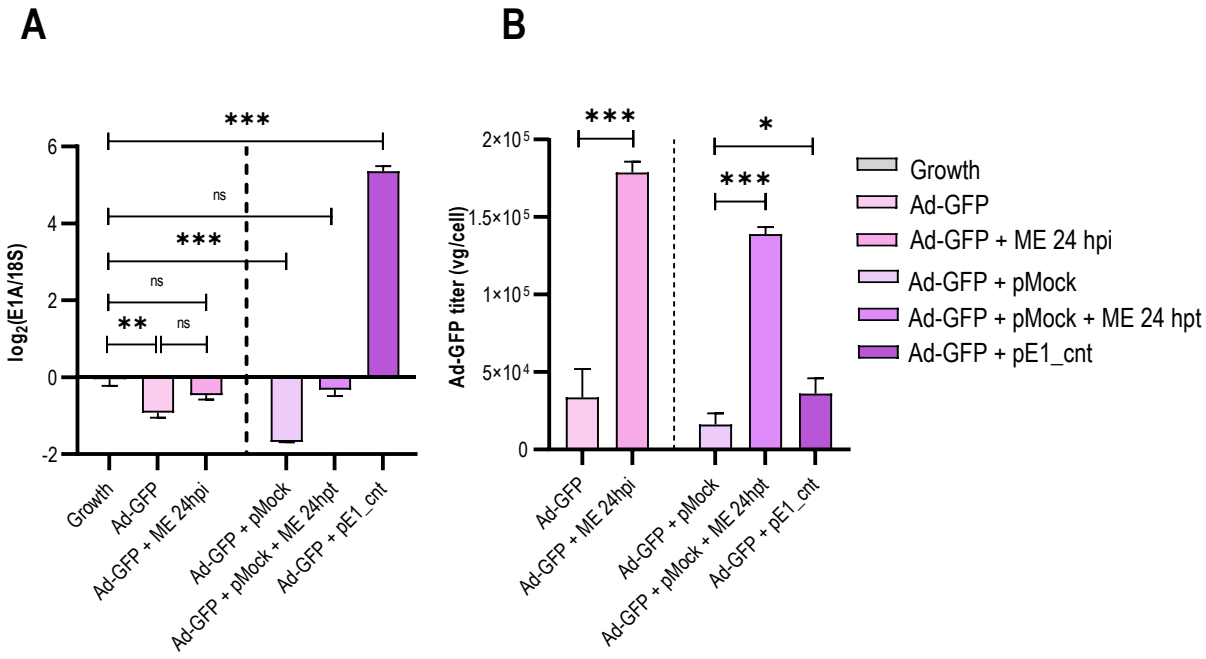


Figure 10. Characterization of E1A preservation and supplementation strategies on Ad-GFP production. A. E1A mRNA expression levels following different protocols at harvest (48 hpi/hpt). **B.** Ad-GFP titer at cell lysate at harvest (48 hpi/hpt). Ad-GFP were produced in HEK293 cells by infection at MOI 10. ME was performed 24 hpi or hpt. Transfections with pMock or pE1_cnt plasmids were performed 4 hpi. Biological triplicates were measured in triplicates. Data was analyzed by ANOVA analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD. Ad-GFP, Adenoviral vector GFP. Hpi, hours post infection. Hpt, hours post transfection. ME, medium exchange. Vg; vector genomes.

DISCUSSION

The incorporation of *E1* adenoviral genes into the *PSG4* gene on human chromosome 19 by Graham *et al.* in the 1970s originated the HEK293 cell line[181,182]. This modification gave birth to the establishment of HEK293 cells as a workhorse for the production of E1-dependent particles in gene therapy and vaccine development context[373]. Notably, the findings reported here indicate that during rAAV production in HEK293 cells, endogenous E1A expression decreases significantly 96 hpt. This decrease is not observed in either proliferating cells or in those transfected with pMock plasmid. Further, preservation strategies of E1A expression within HEK293 cells result in an enhancement of rAAV production. These observations highlight the role of the adenoviral *E1* genes as a bottleneck in the E1-dependent viral particle manufacturing process within the HEK293 cell line, and potentially, E1-transformed cell lines.

This has substantial implications for both transient transfection-based and stable cell line-based rAAV production processes. In the transient transfection context, the preservation of E1A expression provides a strategy to enhance rAAV production as demonstrated here but could also potentially extend the production window beyond 96 hpt, leading to a higher rAAV accumulated yield. In this work, harvest is performed at 72 hpt to ensure comparability of the different production methods.

In the process of stable cell line generation for rAAV production, extensive efforts have been directed towards achieving a tight inducible control of the cytotoxic *rep* gene in HEK293-based cell lines[255,256,258,275]. Nonetheless, the results provided herein indicate that E1A may play a critical role in rAAV production, even if tight control of *rep* gene is achieved. Of note, attempts to create cell lines with inducible E1A expression have encountered several challenges related with the decrease in inducibility of E1A expression over time, and basal levels of integrated AdV *E2A, E4* and *VA RNAs* proved toxic to the cells, leading to cell lines that were unstable and inefficient for rAAV production[257]. This may be explained in the context of the results provided here: E4 ORF or VA RNAs not only negatively impact cell viability, as previously documented in the literature[198,374], but it also have an inhibitory effect on E1A expression. However, the underlying mechanism behind these interactions remains unclear and deserves further investigation. Future studies could delve deeper into E4 ORF, VA RNAs and E1A interplay, aiming to clarify their mechanistic relationship. Nonetheless, these results evidence the relevance of tightly regulating adenoviral helper genes in a producer rAAV cell line, avoiding leaky expression of VA RNAs and E4 ORF.

Two different strategies are proposed to counteract E1A expression decrease during rAAV production: (1) Media exchange during rAAV production process, inducing endogenous E1A expression and (2) pE1_cnt plasmid co-transfection with TT plasmids, providing exogenous E1A supplementation.

The process of replacing spent media serves multiple purposes in the rAAV production context. Metabolically, it counteracts the observed decrease in E1A during rAAV production. In terms of production outcomes, it enhances rAAV production and promotes the rAAV egress. Interestingly, the findings reported here indicate that the impact of media exchange on rAAV egress varies depending on the serotype, aligning with prior studies that have documented serotype-dependent differences in rAAV release properties[152]. In addition, it has been reported that virus concentration in the cell media depends on each rAAV serotype affinity for specific receptors, particularly cellular heparan sulfate proteoglycan (HSPG). Since HSPG acts as a primary receptor for AAV2, this serotype is predominantly cell pellet bound[152]. Therefore, the strategy of harvesting the cell culture media during production may reduce the potential for secreted rAAVs to re-bind and enter the producer cells, maintaining them at the extracellular fraction[156]. Of note, the strategy of replacing exhausted media can be efficiently implemented into continuous manufacturing bioprocesses using cell retention devices. This continuous mode of operation offers dual benefits: it ensures a consistent supply of fresh nutrients and the efficient removal of waste and toxic by-products, thus maintaining optimal culture homeostasis. Additionally, it facilitates continuous harvesting of rAAV from the supernatant, which eliminates the need of lysing the producer cells. This aspect is particularly beneficial as it simplifies the subsequent purification steps, a substantial advantage for large-scale production. Such perfusion system has been implemented for transient transfection rAAV production[210] and rAAV stable producer cell lines[339].

E1A increase by E1 overexpression using metabolic engineering did not altered viral egress but demonstrated a greater enhancement on accumulated rAAV titer compared to media exchange for all tested serotypes, both at the cell lysate and supernatant fractions. The underlying mechanism by which E1 enhances rAAV production is likely multifactorial. E1A is known for its role in creating an optimal environment for viral replication by promoting cell cycle progression, overcoming growth arrest, and inducing cells to enter S-phase[375,376]. E1A stimulates entry into S phase by cell cycle deregulation by binding to the retinoblastoma protein (Rb), and the transcriptional coactivator p300/CBP which is also involved in the process of E1A-induced stimulation of E2F[324,377,378] . Free E2F also binds and induces viral genes such as E2A, stimulating both cellular and viral genes needed for viral DNA replication. Coherently, in **Chapter one**, the transition from the G1 to S phase

of the cell cycle was hindered in TT, while in the EGE protocol the downregulation was prevented. Additional work would be required to analyze if the E1A supplementation methods shown here also involve cell cycle transition alterations.

Controversially, functions of E1A that are required for efficient rAAV production also trigger cell death by apoptosis, both by a p-53 dependent and p-53 independent mechanisms[111]. Considering E1A pro-apoptotic properties, both E1B 19K and 55K co-expressed in the adenoviral genome exhibits anti-apoptotic activities, presumably to counteract the effect of E1A and sustain productive infection and transformation of hosts cells[112–114] In the proteomic analysis presented in **Chapter one**, the E1B 55K protein expression decreased over TT rAAV production, and while it eventually recovered in standard production method, it was consistently downregulated in the EGE protocol. Similarly, through the metabolic engineering approach that is presented here, both E1A and E1B functions are provided in pE1_cnt plasmid, aiming to contribute to rAAV replication and extending the cell's lifespan through the anti-apoptotic properties of E1B. Contrarily, Gu *et al*. reported that greater enhancement of rAAV titer was achieved when only *E1A* gene was added by transient transfection compared to when E1 complete ORF was provided[379]. Overall, further data would be required to analyze how E1A supplementation strategies increase rAAV production and how apoptotic processes are altered during the process.

In this study, it is reported that both ME and metabolic engineering strategies, impacting E1A expression into HEK293 cells, influence its physiological characteristics and the way cells respond to rAAV production, achieving different yields. One possible explanation for the difference in rAAV production outcome by the two explored methodologies of E1A expression preservation may be the difference in E1A mRNA localization. Pelka *et al*., in 2021 showed that E1A mRNA of AdV infection in *cis*, colocalized to virus replication centers whereas E1A RNA intrinsically expressed by HEK293 did not[380]. This difference in subnuclear localization also caused an altered splicing pattern. Although this comparison was made with adenovirus carrying E1A in *cis*, we hypothesize that increasing E1A levels by exogenous plasmid co-transfection or by simply restoring endogenous E1A levels by HEK293 itself as in ME, may be governed by a similar principle. Overall, further investigation is needed to determine if differential subnuclear localization and metabolic modulations exist within the cell between exogenously introduced or endogenously produced E1A.

In summary, the findings reported here underscore the role of endogenous E1A expression as a limiting factor in rAAV production in HEK293 cells. The strategies proposed to preserve or increase E1A expression levels in the producer cells during rAAV production represent a step forward towards the optimization of large-scale manufacturing platform for rAAV production processes.

CONCLUSIONS

This study identifies endogenous *E1A* gene expression in HEK293 cells as a significant bottleneck in rAAV production. Specifically, the *VA RNAs* and *E4* adenoviral helper genes within the pHelper plasmid contribute to the reduction in E1A expression during rAAV production by standard triple transfection. Media Exchange (ME) is identified as a key modulator of endogenous E1A expression, effectively preserving it during rAAV production. Additionally, E1A overexpression through metabolic engineering not only restores but also overexpresses E1A levels. E1A expression modulation by both strategies enhances the production of rAAV in both supernatant and cell lysate fractions, with E1A increase by metabolic engineering outperforming ME in rAAV titers. Interestingly, ME is identified as the critical step in the Extended Gene Expression (EGE) method that promotes rAAV egress. These findings pave the way for developing more productive viral production methods by Transient Gene Expression (TGE) or Stable Gene Expression (SGE) approaches by the effective modulation of *E1A* gene expression.

PART II. Proof of concept of an inducible system towards stable rAAV cell line generation

CHAPTER FOUR

Rational design of optimal inducible Rep construct for the generation of a stable rAAV packaging cell line in HEK293 cells

INTRODUCTION

To date, several strategies have been developed for recombinant adeno-associated virus (rAAV) production[99,106,216]. Currently, transient co-transfection of HEK293 cells is the approach of choice for rAAV production. This strategy, based on transient gene expression (TGE), can yield sufficient vector amounts for therapeutic transgene screening, clinical trials and even final production scale[293]. While it offers versatility in terms of the final product, it also brings significant challenges in large-scale manufacturing. Multiple plasmid transfection results in high cost of goods and high variability of the process, challenging the robustness during process scale-up[194]. Additionally, the presence of prokaryotic sequences in plasmid DNA may rise safety concerns for reverse packaging[149]. Moreover, although TGE is a rapid and straightforward method to express the product of interest, plasmid DNA is not integrated, thus, its expression is limited in time.

In contrast, producer cell lines achieving stable gene expression (SGE) are preferred for biotherapeutics production[194]. In large scale bioprocessing, SGE strategy is preferred over TGE, as it eliminates the need of DNA and transfection reagents, providing a more economic and simple production system, with less batch-to-batch variability[381]. This approach involves the direct integration of the genes encoding the protein product into the host cell genome, leading to constitutive expression[382]. However, the process of generation, screening and selection of the final clones is time-consuming, and may take up to 12 months[383,384].

In the context of rAAV production, the SGE strategy has been investigated as a promising alternative to TGE systems by generating packaging or producer cell lines that stably encode the essential genes required for rAAV production within their genome. Packaging cell lines represent a first step towards the development of producer cell lines as they only contain the *rep* and *cap* genes of the serotype of interest. Thus, still require additional functions for rAAV production, such as helper virus genes which are typically supplied by infection, which is associated with significant operational challenges such as helper virus removal during purification, and is highly undesirable considering safety[188,247]. In contrast, producer cell lines containing in their genome all components required for rAAV production, including: the GOI, *rep* and c*ap* genes, and the adenoviral *helper* genes are desirable, although few successful attempts have been reported to date, showing low viral titers[248,249].

Several challenges still need to be addressed for the development of rAAV packaging or producer cell lines[246]. The main limitation is the cytostatic and cytotoxic properties of viral proteins required

for rAAV production, such as Rep from AAV[58,197] and Helper proteins E2A and E4 from AdV[257]. To address these challenges, the choice of the host cell is key to determine the approach to overcome the toxicity[250–254]. Using HEK293 for generating packaging cell lines for rAAV production, containing *rep-cap* genes integrated is particularly challenging as it constitutively express *E1A* and *E1B* genes from AdV[181] leading to constant E1A-mediated transactivation of the AAV p5 promoter that control the expression of Rep proteins[58,104,197] (**Figure 1A, B**). Large Rep proteins (Rep78 and Rep68) have a strong antiproliferative effect. Particularly, Rep78 induces cytotoxic and cytostatic effects in host cells, leading to cell death through a p53-independent apoptosis mechanism mediated by caspase-3[58]. This process is reported to be cell cycle-specific, with Rep78 accumulation observed in cells in the G1 phase of the cell cycle. The cause of cell death has been attributed to DNA binding and ATPase/helicase activities[58]. In addition, Rep78 and Rep68 block cell cycle progression differently due to the differential zinc finger domain (ZFD). While Rep78 induces S-phase arrest with the accumulation of hyperphosphorylated retinoblastoma protein (pRb), Rep68 induces arrest at both G1 and G2 phases[57,59].

Therefore, considering Rep toxicity as the main limitation towards a HEK293-based packaging cell line generation, several strategies to tightly control its expression have been developed to date[246]. Remarkably, while inducible promoters have been used to replace p5 promoter and control large Rep toxic protein expression[197,258,385], leaky expression of small Rep proteins, expressed from p19 promoter still lead to cell growth limitations (represented in **Figure 1A, B**)[48,197,252,255,259].

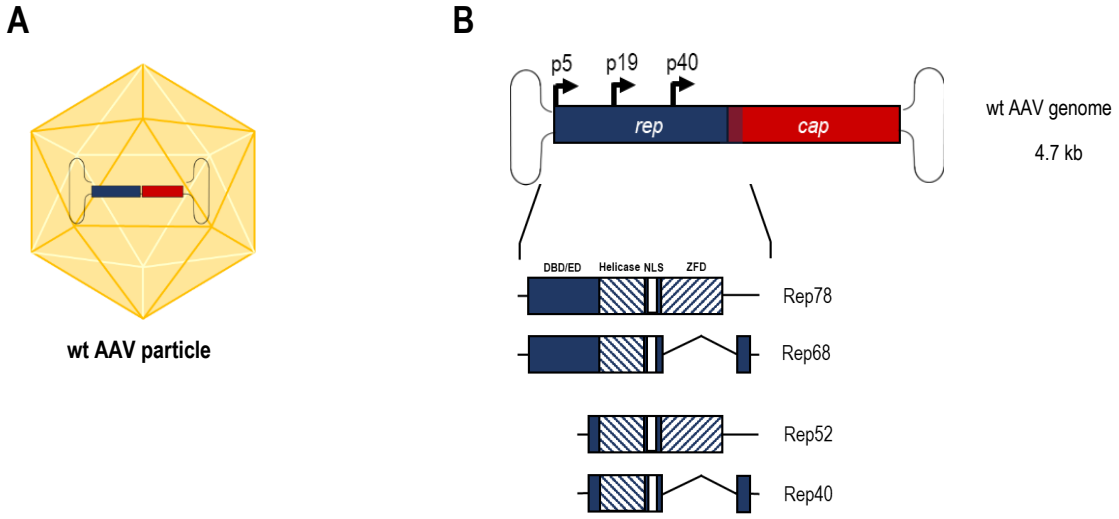


Figure 1. Schematic representation of AAV genome. A. wt AAV virion scheme. **B.** Schematic representation of the wt AAV genome organization. Relative positions of *rep* and *cap* genes are depicted. The 145-nt ITRs, which serve as viral origins of replication are indicated at both ends. Rep78 is shown at the top with the different functional domains. DBD, DNA-binding domain; ED, endonuclease domain; NLS, nuclear localization site; Wt, wild-type; ZFD, zinc finger domain.

Two different strategies have been proposed so far to address the challenges by Rep expression in the development of HEK293-based packaging cell lines. One approach involves regulating all Rep proteins using the "dual splicing switch" strategy based on the Cre-lox system[256,260]. The other method is based on the individual expression of only one large and one small Rep protein from separate constructs, controlling each expression by an inducible promoter[247,275]. However, these approaches represent incomplete systems, and to produce rAAV, additional transfection with the pGOI plasmid, followed by infection with AdV or transfection with helper functions is needed. Additionally, inducible systems also require the induction of the rep element, present in the cell genome[247,275]. Although packaging cell lines represent a first step towards complete producer cells, low titers and suboptimal packaging efficiency has been reported[247]. Additionally, the optimal combination of *rep* genetic elements leading to a maximal and proportional dose-response between the inducer and rAAV production has not been addressed, which is a key feature for the development of packaging cell lines based on an inducible system[49,386].

In the current study, as a first steps towards the primary objective of generating a HEK293-based stable cell line for rAAV production, two goals are aimed: (1) to elucidate the minimal Rep proteins required to support maximal rAAV production under an inducible system and (2) minimize the potential toxicity of the rep element towards its integration into HEK293 cells.

To this end, a stepwise rational genetic engineering of the *rep* element is performed. This involves separating the expression of the four Rep proteins into different plasmids with mutations that lead to its individual expression. Subsequently, a tight tetracycline-responsive promoter is used to control their expression under the Advanced Tet-On® inducible system[387]. By means of a Design of Experiments (DoE)-driven approach, the optimal combination of large and small Rep proteins required for maximal rAAV production is identified. This process also includes the assessment of critical factors influencing the rep inducible system. Furthermore, the impact of mutations introduced in the large Rep construct, as well as the effects of intron maintenance and small Rep transcript expression on rAAV titer and toxicity potential are evaluated.

The gene engineering of the *rep* gene shown here represents a significant step towards the successful generation of a packaging cell line for rAAV production in cells constitutively expressing the E1A protein, leading the groundwork for more robust rAAV production systems potentially broadening the scope of gene therapy applications.

MATERIALS AND METHODS

Mammalian cell line and culture conditions

The HEK293 cell line (HEK293SF-3F6) used is a suspension-adapted serum-free cell line kindly provided by Dr. Amine Kamen from the National Research Council of Canada (NRC, Montreal, Canada). A serum-free, suspension-adapted A549 cell line was also used kindly provided by Theriva™ Biologics (Parets del Vallès, Spain).

HEK293 cells were cultured in HyCell™ TransFx-H medium (GE Healthcare, Chicago, IL, USA) supplemented with 4mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA) and 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, ThermoFisher Scientific, San Jose, CA, USA).

The medium used for suspension A549 cell culture was ExCell® CD HEK293 Viral Vector media (Merck, Darmstadt, Germany) supplemented with 6 mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA), 1x N-2 MAX Media Supplement (Bio-Techne, Minneapolis, MN, USA) and 0.6% Kolliphor (Gibco, ThermoFisher Scientific, San Jose, CA, USA).

When required according to planned experiments, culture media was supplemented with doxycycline (Sigma-Aldrich, San Luis, MO, USA) dissolved in water.

Suspension cell cultures were maintained in exponential growth phase in disposable polycarbonate 125 mL baffled flasks with vent cap for A549 cell line or non-baffled flasks for HEK293 cells (Corning®, Tewksbury, NY, USA) at 37ºC, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Cell concentration and viability were determined using the NucleoCounter® NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions. Cultures were observed under the Olympus CKX41 optical inverted microscope (Olympus, Tokyo, Japan) to qualitatively analyze cell growth and viability.

Plasmid design and generation

Plasmids used for standard triple transfection were kindly provided by Dr. Fàtima Bosch from Center for Animal Biotechnology and Gene Therapy (CBATEG), at Universitat Autònoma de Barcelona (UAB, Bellaterra, Spain). These included plasmids pXX6, pGG2_CMV_GFP (referred to as pGFP_ITR hereafter), and pRep2Cap1. The plasmid pRep2Cap1 encodes the Rep proteins from AAV2 and capsid proteins from AAV1. The pGFP_ITR plasmid encodes for the enhanced GFP gene under the control of the cytomegalovirus (CMV) promoter and is flanked by two inverted

terminal repeats (ITRs) to enable its packaging into rAAV capsids. The pXX6 plasmids codifies for the adenovirus elements necessary for rAAV replication and packaging[99].

For transfection control, pMock plasmid was used. This plasmid, devoid of any mammalian promoter or coding DNA sequence (CDS) was engineered by ligating the pGag::eGFP backbone[350,351].

For the separate expression of Rep and Cap genes, pRep_wt and pCMV_CAP1 plasmids were designed and constructed. To generate the pRep_wt plasmid, Rep open reading frame (ORF) from AAV2 was PCR amplified using Q5 High-Fidelity DNA polymerase (New England Biolabs (NEB), Ipswich, MA, USA) from the pRep2Cap9 plasmid using phosphorylated primers following manufacturer's instructions. Resulting fragments were purified and ligated using T4 DNA ligase (NEB, Ipswich, MA, USA). To construct pCMV_CAP1 plasmid, the VP1 ORF from pRep2Cap1 plasmid was amplified using Q5® High-Fidelity DNA polymerase (NEB, Ipswich, MA, USA) following manufacturer's instructions. Concurrently, backbone region from pGag_GFP plasmid's backbone was amplified maintaining CMV enhancer and promoter region to drive capsid gene expression. Primers were designed to generate overlaps of 20 nt of homology between fragments. Two-fragment assembly was performed using the NEBuilder® HiFi DNA Assembly mix (NEB, Ipswich, MA, USA) for 1 hour at 50ºC using a molar ratio of 1:2. Primers used for cloning are shown in **Table 1.**

Inducible expression following the Tet-On® system requires two mammalian expression vectors. For the construction of the plasmid coding for constitutive expression of the reverse transactivator rtTA Advanced® (Clontech, Takara, San Jose, CA, USA), the CDS containing rtTA-Advanced ($rtTA2^{S}$-M2)[388] and WPRE were obtained as a gBlock at IDT (Integrated DNA Technologies, Leuven, Belgium) flanked by regions with sequence complementarity to a backbone containing CMV promoter and SV40 polyA signal (from pGagGFP plasmid). Both the gBlock coding for rtTA and pGagGFP were PCR amplified using primers shown in **Table 1**. Products were checked by gel electrophoresis and then incubated with Hi-Fi Master Mix 2x (NEB, Ipswich, MA, USA) following manufacturer's instructions.

Table 1. List of primers used for cloning. Sequences added to generate sticky ends or homologies for HiFi DNA assembly are underlined. Introduced mutations are shown in lowercase. Primers were designed avoiding possibilities of self-dimerization and primer-dimer formation.

Name	Objective	Sequence (5'-3')
Fwd_Rep2_wt	Amplify Rep78 ORF from pRep2Cap9 plasmid	TGTTTCCTTGCGTATTGGGC
Rev_Rep2_wt		ACCTGGTTTAAGTCATTTATTGTTC
Fwd_CMV_CAP1	Amplify Cap1 ORF from pRep2Cap1 plasmid	GTGAACCGTCAGATCCGCTAATGGCTGCCG ATGGTTATCT
Rev_CMV_CAP1		TATGTAACGCGGAACTCCATCTCACTGGCC GTCGTTTTAC
Fwd_CMVenh	Amplify CMV enhancer and promoter region from pGag_GFP plasmid	ATGGAGTTCCGCGTTACATA
Rev_CMVenh		TAGCGGATCTGACGGTTCAC
Fwd_GFP_ind_mod1_NotI	PCR amplification of pGFP_ind_mod1 backbone	CCTGACGCGGCCGCCGCTCACTGACTCGC TGCGC
Rev_GFP_ind_mod1_EcoRI		TTGACAGGGAATTCGGCGACCGGTGGATCG AAGC
Fwd_GagGFP_rtTA	PCR amplification of pGagGFP backbone for cloning into rtTA gBlock	GGCCGCGACTCTAGATCATAATC
Rev_ GagGFP_rtTA		TAGCGGATCTGACGGTTCACT
Rev_rtTA_gblock_ NotI	PCR amplification of rtTA gBlock	GGCTGATTATGATCTAGAGTCGC
Fwd_rtTa_gblock		CACGCTGTTTTGACCTCCAT
Fwd_Rep40	PCR amplification of Rep small Strings	CTCTAGTCAGCTGACGCCGG
Rev_Rep40_Eco43I		CAGTGATATCATAAGGCGCCAAGTTTGAGC TTCCA
Fwd_68	PCR amplification of Rep large Strings	CTAGTCAGCTGACGCTCCCC
Rev_68		ATTATGATCCGCGGATCCGCG
Fwd_mutTATA	PCR mutagenesis of pRep78 to revert p19 mutations	ATGGAACAGTAtTTAAGCGCCTGTTTGAATC TCACGGAGCGTAAAC
Rev_mutTATA		CAAACAGGCGCTTAAaTACTGTTCCATaTTA GTCC
Fwd_mutS	PCR mutagenesis of pRep78 to revert splicing site mutation	gTACCAAAACAAATGTTCTCGTCACG
Rev_mutS		CTGTCTGCGTAGTTGATCGA

To construct Rep expression plasmids for Tet-On® inducible system, Rep ORFs were rationally modified as specified in **Table 2** based on AAV2_ AF043303.1. For Rep78 and Rep52 ORFs, splicing sites were mutated (from CAGGTACCA to CAGATACCA) to maintain the intron sequence[247]. For Rep68 and Rep40 ORFs, intron was deleted. Mutations were introduced in the internal AAV p19 promoter in Rep78 and Rep68. Specifically, TAATAT to TAACAT from TATA box -35 and TATTTA to TACTTA from TATA box -20 mutations were performed to avoid Rep small transcript expression, similarly as in Jalšić *et al.*[247]. Additionally, p40 TATA box -19 was modified in all ORFs to suppress any potential additional ORF from GATATA to GACATA. A mutation for later cloning was introduced at Rep small ORFs at 1048 nt position from C to T. All mutations were designed avoiding affecting the amino acid sequence of wild-type (wt) Rep proteins[247].

Table 2. Summary of mutations and modifications of the newly rationally designed Rep ORFs.

Rationally designed Rep ORF[a]	Nucleotides[b]	p19 TATA mutations at -35 and -20 (830 and 843 nt respectively)	Splicing site mutation (at 1907 nt) / deletion
Rep78	321 to 2252 nt	TAATAT to TAACAT from TATA box -35 and TATTTA to TACTTA from TATA box -20	from CAGGTACCA to CAGATACCA
Rep68	321 to 1906 and 2228 to 2252 nt	TAATAT to TAACAT from TATA box -35 and TATTTA to TACTTA from TATA box -20	Deletion of intron (nt 1907-227)
Rep52	993 to 2252 nt	NA[c]	from CAGGTACCA to CAGATACCA
Rep40	993 to 1906 and 2228 to 2252 nt	NA[c]	Deletion of intron (nt 1907-227)

[a] All Rep ORF contained p40 TATA -30 mutation at 1822 nt.
[b] Numbers based on AAV2 AF043303.1.
[c] NA: non aplicable.

To construct the tetracycline-responsive element (TRE) upstream of the designed Rep CDS, the described Rep78, Rep68, Rep52 and Rep40 ORFs were synthetized as DNA String fragments (GeneArt, ThermoFisher Scientific, San Jose, CA, USA) with the minimal CMV promoter sequence (CMVmin) at 5' end. Additional nucleotides were incorporated at both ends of String fragments for subsequent cloning with pTRE_tight_BI plasmid. pTRE_tight_BI plasmid was synthesized (GeneArt, ThermoFisher Scientific, San Jose, CA, USA) and designed to incorporate suitable enzyme restriction sites at both ends of TRE_tight bidirectional promoter.

Finally, each individual Rep ORF were cloned into pTRE_tight_BI. Rep large ORF were PCR amplified using primers shown in **Table 1** and cut with XmaI and BamHI (NEB, Ipswich, MA, USA) and inserted into the pTRE_tight_BI using the same restriction enzymes, positioned at 3' of the

TRE in pTRE_tight_BI plasmid. Rep small ORF were PCR amplified using primers shown in **Table 1** and cut with Eco32I and KpnI (NEB, Ipswich, MA, USA). The isolated DNA was inserted into pTRE_tight_BI vector cut with EcoRV and KpnI (NEB, Ipswich, MA, USA). This resulted in Rep small ORFs positioned 5' of the TRE, at inverse orientation.

To reverse pRep78 mutations, PCR point mutagenesis protocol described by Zheng *et al.* was employed[389]. Briefly, the plasmid pRep78 was amplified using two complementary oligonucleotides (**Table 1**). After amplification, the PCR product was digested with DpnI (NEB, Ipswich, MA, USA) for 16 hours at 37ºC, a methylation-dependent endonuclease that only digests the non-PCR amplified plasmid. First, splicing site mutation was reverted and pRep78mutS plasmid was generated. This plasmid was used for reverting TATA mutations at p19 promoter using a similar protocol as previously described, generating pRep78mutTS. Finally, pRep78mutT plasmid was constructed.

Plasmids were confirmed by Sanger sequencing (Servei de Genòmica, UAB, Bellaterra, Spain) and amplified in a chemicompetent *E. coli* DH5α strain produced *in house* growing in LB medium (Conda, Madrid, Spain) supplemented with 10 µg/ml kanamycin (Sigma-Aldrich, St. Louis, MO, USA) or 100 µg/ml ampicillin (Sigma-Aldrich, St. Louis, MO, USA) depending on the *E. coli* antibiotic resistance present on each plasmid. Plasmid purification was carried out using the Endofree Plasmid MIDI kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Quantity of isolated DNA was assessed using Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, San Jose, CA, USA).

Standard transient transfection protocol and rAAV production

Transient transfection of HEK293 suspension cells was performed at a cell density of 2 x 10^6 cells/mL using 25-kDa linear polyethilenimine (PEI) (PEIPro®, Polyplus-transfection SA, Illkirch-Graffenstaden, France). Transfections were carried out using a final DNA concentration of 1 µg/mL with a DNA:PEI ratio of 1:2 (w/w).

Transient transfection in A549 cell line was performed at 2.8 x 10^6 cells/mL using PEI as transfection reagent. Total DNA concentration used was 3.6 µg/mL and a DNA:PEI ratio of 1:2 (w/w) according to the transient transfection optimization performed in **Chapter five.**

For transient transfection, plasmid DNA was diluted in supplemented HyCell™ media in HEK293 or non-supplemented Ex-Cell™ media in A549 cells (10% of the culture volume to be transfected) and vortexed for 10 seconds. Then, PEI was added and vortexed three times for 3 seconds,

incubated for 15 minutes at room temperature and then added to the cell culture. Transient transfections were performed in six-well plate in suspension in 2 mL total culture volume.

For rAAV production, transient transfection was performed in HEK293 cells using equimolar plasmid ratio unless stated otherwise for experimental requirements and normalized to 1 µg/ml using pMock when necessary.

For rAAV harvest, cell culture samples were taken at 72 hpt. For cell supernatant collection, cell culture sample was centrifuged at 1000 x g for 10 minutes. For cell lysate collection, whole cell culture was treated with 1x cell lysis solution (500 mM Tris, 1 % Triton X-100 V/V, 20 mM $MgCl_2$ at pH 7.5) and incubated at 37ºC, 130 rpm, for 1 hour. Lysate mixture was centrifuged at 4000 x g for 20 minutes and supernatant was collected as lysate fraction. All fractions were stored at -80ºC until analysis was performed.

Optimization of quadruple transfection for rAAV production by design of experiments (DoE)

Quadruple transfection was optimized for rAAV production to evaluate the feasibility of Rep-Cap split plasmid design. A four-factor, three level Box-Behnken experimental design was selected to determine the optimal concentration of the four transfected plasmids that maximizes rAAV titer, namely: pHelper, pGFP_ITR, pRep_wt and pCMV_CAP1. Each of these variables were screened across three different levels shown in **Table 3**. rAAV production response was converted to logarithmic scale to normalize the variability of the residuals. Box-Behnken experimental results were fitted to a second-order polynomial equation by linear regression analysis for each response (**Equation 1**):

$$\mathrm{Y} = \beta_0 + \sum_{i=1}^{k} \beta_i \cdot X_i + \sum_{i=1}^{k} \beta_{ii} \cdot X_i^2 + \sum_{i=1}^{k} \sum_{j>1}^{k} \beta_{ij} \cdot X_i \cdot X_j \qquad (1)$$

where Y is the response variable under consideration (rAAV titer (vg/cell)), β_0 is the model intercept term, β_i is the linear coefficient, β_{ii} is the quadratic coefficient, β_{ij} is the interaction coefficient, and the coded factors X_i and X_j are the independent variables of study. This equation facilitates the exploration of *k* factors at three levels. In this context, k=4, corresponding to the four independent variables considered in this study. This equation was used to predict the optimal values of the independent variables using the Design-Expert® Software (Stat-Ease, Inc, Minneapolis, MN, USA).

Mixed-level factorial design to determine the optimal combination for rAAV production using inducible Tet-On® system

Considering that rAAV titer under the inducible Tet-On® system is affected by initial Rep copy numbers of plasmid and inductor dose, a DoE was conducted varying these factors to select the optimal Rep construct combination that yields higher rAAV titer. Specifically, a mixed-level factorial design was selected to determine the main effects and interactions of different factors on rAAV titer. Four different variables were evaluated at two or three different levels to identify optimal Rep inducible constructs combination in relation to the evaluated factors. The variables screened are shown in **Table 4**.

First, doxycycline concentration was varied at two levels: low (coded as +) and high (++), corresponding to 0.05 and 1 µg/mL of doxycycline respectively. Then, Rep total quantity of added plasmids was also varied at two levels, low of 1.40 x 10^8 and high of 3.50 x 10^8 copies/mL. This were set in the proportional dose-response range according to previous experiments. Finally, optimal ratio needed of Rep large and small constructs was determined independently. Large rep constructs were varied from pRep78 and pRep68. Rep small transcripts were varied from pRep52 and pRep40. In these variables, a low level of no Rep plasmid addition (0) was considered for both factors to determine if no addition can be considered for each Rep transcript combination. Rep large and small transcripts concentration varied according to Rep total quantity in each condition.

Response was measured as rAAV specific titer per viable cell density (VCD) at harvest (vg/cell). Combinations of the model in which both Rep large and small transcripts correspond to “0” level were depicted as NA, as no rAAV production results from this combination. Response was converted to logarithmic scale to normalize the variability of the residuals. Mixed-level factorial design results were fitted to a polynomial equation by linear regression analysis to determine the main effects and interaction effects (**Equation 2**):

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i \cdot X_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} \cdot X_i \cdot X_j \qquad (2)$$

where Y is the response variable under consideration (rAAV titer (vg/cell)), β_0 is the model intercept term, β_i is the linear coefficient for the main effect of factor X_i, β_{ii} is the coefficient for the interaction effect between factors X_i and X_j , which are the independent variables of study. This equation facilitates the exploration of *k* factors at different levels. In this context, k=4, corresponding to the four independent variables considered in this study. This equation was used to predict the

optimal combination of each level of the independent variables using the Design-Expert® Software (Stat-Ease, Inc, Minneapolis, MN, USA).

DoE statistical analysis

Statistical analysis and three-dimensional graphs of the different DoEs facilitating model interpretation was performed using Design-Expert® Software (Stat-Ease, Inc, Minneapolis, MN, USA). The quality of the fit of the model equation is expressed by the coefficient R^2. The lack-of-fit (LOF) test was used to evaluate differences between experimental and pure error of the fitted equations. The overall significance of the model was determined by analysis of variance (ANOVA) *F*-test, whereas the significance of each coefficient was determined by the corresponding *t*-test. *p*-values of 0.05 and 0.1 were considered as statistically significant with 95% and 90% confidence, respectively.

Quantitative polymerase chain reaction (qPCR) for rAAV quantification

For the determination of physical vector genome titer of rAAV (vg), 5 µL of the cell lysate fraction or cell supernatant fraction from each sample was added to DNase I (0.066 mg/mL, 10104159001, Roche, Basel, Switzerland) prepared in DNase I buffer (0.16 mM $CaCl_2$, 6.6 mM $MgCl_2$, 17.3 mM Tris-HCl, 0.066 % Pluronic F68, pH 8.0). This mixture was incubated at 37 ºC for 16 hours. Then, 4 µL of 30 mM EDTA (Invitrogen, Carlsbad, CA, USA), in 10 mM Tris-HCl, pH 8.0 was added to each tube and incubated at 70ºC for 30 minutes to inactivate DNase I. After, 5 µL of Proteinase K (EO0491, ThermoFisher Scientific, San Jose, CA, USA) were added, incubated at 55ºC for 2 hours and inactivated at 95ºC for 15 minutes. The quantification of viral genomes was conducted using quantitative PCR (qPCR). This reaction was set up using iTaq Universal SYBR Green Supermix (BioRad, Hercules, CA, USA) and a set of primers specific to a region within the eGFP gene (**Table 5**). Thermocycling profile was based on manufacturer's recommendations in a CFX384 Touch Real-Time PCR Detection System (**Table 6**). For the establishment of a standard curve, a gBlock from IDT (Integrated DNA Technologies IDT-DNA, Leuven, Belgium) containing the targeted eGFP region of known concentration was used (sequence shown in **Table 5**). A serial dilution of this standard, ranging from 10^8 to 10^2 copies/µL) was prepared and amplified by qPCR. The concentration of AAV samples was determined by interpolating in a standard curve. Additionally, an internal rAAV1-CMV-GFP reference was amplified alongside the samples in all the quantifications.

Table 5. Primers and gBlock sequences used for qPCR quantification of rAAV titer.

Name	Sequence (5'-3')
Fwd_GFP_qPCR	AGCTGGACGGCGACGTAA
Rev_GFP_qPCR	GTGCAGATGAACTTCAGGGTCA
GFP gBlock	AGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC CACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGA C

Table 6. Conditions of cycling temperatures and time for qPCR and RT-qPCR**.**

Reaction	Cycles (temperature and time)
qPCR for rAAV quantification and RT- qPCR	95ºC, 3 min + 40 x [95ºC, 10 sec + 60°C, 30 sec] + 95ºC, 10 sec + 65 to 95°C, 0.5ºC/5 sec + 4°C hold

Quantitative reverse transcription PCR (RT-qPCR)

Total RNA extraction from cell pellets, stored at -80ºC was performed using the Promega Maxwell RSC simplyRNA Tissue Kit (Cat No. AS1340) according to the manufacturer's protocol. The quantity of isolated RNA was assessed using Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, San Jose, CA, USA). cDNA was synthesized using 0.75 µg of total RNA per sample with the iScript cDNA Synthesis Kit (BioRad, Hercules, CA, USA).

Quantitative reverse transcription PCR (RT-qPCR) assays were performed using the iTaq Universal SYBR Green Supermix (BioRad, Hercules, CA, USA) on a CFX384 Touch Real-Time PCR Detection System. Each qPCR reaction mixture contained 2 µl of 3x diluted cDNA and a primer concentration of 500 nM. At the end of the PCR, a melting curve was recorded from 65 to 95ºC. The qPCR assay was performed in triplicate for all samples.

Relative gene expression levels were quantified using the $2^{-\Delta\Delta Ct}$ method[365]. Three set of primers were designed to target different regions of the Rep ORF. Primer set 1 was tailored to amplify large Rep transcripts. Primer set 2 targeted unspliced Rep ORF species, and primer set 3 was specific for spliced species. For normalization purposes, primers targeting the 18S ribosomal gene were employed as the housekeeping control. The sequences for all primers, including those for the housekeeping gene, are provided in **Table 7.**

Table 7. List of primers for RT-qPCR.

Name	Sequence (5'-3')
Fwd_Rep_F1	AGCGACCTTGACGAGCATCT
Rev_Rep_R1	GCAACTCCCATTCCTTCTCG
Fwd_Rep_F2	CAACCCGTTTCTGTCGTCAA
Rev _Rep_R2	TCTGGCACCTTTCCCATGAT
Fwd_Rep_F3	TCAAAAAGGGTGGAGCCAAG
Rev _Rep_R3	CGAGCCAATCTGTCTGCGTA
Fwd_18S	TCAAGAACGAAAGTCGGAGG
Rev_18S	GGACATCTAAGGGCATCACA

The calculation of relative gene expression involved several steps. Briefly, the Ct value of each target gene was subtracted from the mean Ct value of the housekeeping reference gene in each sample to obtain ΔCt. For large total transcript expression, the ΔΔCt was calculated by subtracting the ΔCt of each target gene to the ΔCt of the reference control sample. Fold changes in relative gene expression were calculated as $\log_2(2^{-\Delta\Delta Ct})$. For the Rep intron retention index and intron splicing index calculation (**Equation 3** and **4**, respectively), the ΔCt was used to calculate the $\log_2(2^{-\Delta Ct})$ and normalized with the $\log_2(2^{-\Delta Ct})$ of primer set 1. All conditions were performed in triplicate. Ct values higher than 30 were classified as non-detectable (ND) and were discarded.

The intron retention index was measured using the **Equation 3**:

$$\text{Intron retention index} = \frac{\log_2\left(2^{-\Delta Ct}\right) \text{ of primer set 2}}{\log_2\left(2^{-\Delta Ct}\right) \text{ of primer set 1}} \quad \textbf{(3)}$$

The intron splicing index was measured using the **Equation 4**:

$$\text{Intron retention index} = \frac{\log_2\left(2^{-\Delta Ct}\right) \text{ of primer set 3}}{\log_2\left(2^{-\Delta Ct}\right) \text{ of primer set 1}} \quad \textbf{(4)}$$

Cell cycle analysis

Freshly harvested cells were pelleted by centrifugation, rinsed, re-suspended in PBS and normalized to 1 x 10^6 cells in 500 µL. Cells were fixed with cold ethanol (70%) for 2 hours at -20ºC. For cell cycle analysis, cells were washed in PBS twice and stained with propidium iodide (PI)

solution (50 µg/mL PI, 200 µg/mL RNase A, 0.1% Triton X-100 (v/v) provided by SCAC, UAB) for 30 minutes at 37ºC prior to data acquisition with CytoFLEX (Beckman Coulter Inc., Brea, CA, USA).

Total capsid determination by ELISA

rAAV1 capsid titters were determined by ELISA (RAAV1 Titration ELISA, cat# PRAAV1, PROGEN Biotechnik GmbH, Heidelberg, Germany) according to the manufacturer's instructions. Samples were serially diluted to fit the standard curve of reference control provided by the kit. The standard curve was generated with the standard kit provided, using a second-order quadratic polynomial curve fit and sample concentrations were automatically interpolated from the standard curve. Absorbances were measured at 450 nm (SPECTROstar Nano, Ortenberg, Germany). Concentration values were averaged across the dilutions that were within the adjusted absorbance range recommended by the vendor and the standard deviation was calculated.

Statistical analysis

All data were represented as mean ± standard deviation. Unless stated, one-way analysis (ANOVA) was used for all statistical analysis in this study. GraphPad was used to conduct statistical analysis. *p* values less than 0.05 were considered statistically significant.

RESULTS

Evaluation of the Rep-Cap Split Packaging System for rAAV production

In the wt AAV genome, capsid gene expression is driven by the p40 promoter, located within the Rep ORF. To enhance control over *rep* and *cap* genes, their expressions were decoupled by separating their ORF into different plasmids, a strategy known as Rep-Cap split packaging system[204,390]. It has been established by previous studies that high capsid expression levels are important for efficient rAAV production[204,275]. Therefore, the entire VP1 coding sequence from AAV1, devoid of p40 promoter, was placed downstream of the strong CMV promoter to construct pCMV_CAP1 plasmid. For *rep* gene expression, the gene sequence based on AAV2 was used[149]. The complete *rep* gene sequence was maintained with its endogenous promoter p5 (**Figure 2A**).

To test the performance of the Rep-Cap split packaging system, the maximal rAAV1 production capacity was assessed through the four-factor, three-level Box-Behnken Design of Experiments (DoE). This design was employed to optimize the concentrations of four plasmids: pHelper, pGFP_ITR, pCMV_CAP1, and pRep_wt for maximal rAAV production.

The experimental design matrix in coded values, response, and statistical analysis are presented in **Table 3**. Data were fitted to a quadratic second-order polynomial equation using a linear regression analysis resulting in the following coded model equation **(Equation 5)**:

$$Y = 2.76 + 0.29X_1 - 0.09X_2 + 0.59X_3 + 0.11X_4 - 0.09X_1X_2 - 0.18X_1X_3 - 14.72X_1X_4 + 0.01X_2X_3 - 0.10X_2X_4 + 0.14X_3X_4 - 0.35X_1^2 - 0.05X_2^2 - 0.23X_3^2 - 0.13X_4^2 \quad \textbf{(5)}$$

where Y is rAAV1 production (in vg/cell); X_1 is the coded value for pHelper, X_2 is the coded value for pGFP_ITR, X_3 is the coded value for pCMV_CAP1 and X_4 is the coded value for pRep_wt. Regression analysis showed that the model fit the data with a coefficient R^2 of 0.9672 indicating consistency with 95% data variability. Further, the statistical significance of the model was confirmed by ANOVA analysis. The Fisher's F-test associated *p* value of <0.0001 indicated that the model was significant. The lack of fit parameter was non-significant (*p* value 0.1353). Values of the tested plasmids pHelper, pGFP_ITR, pCMV_Cap1 and pRep_wt were also significant (see **Table 3**).

Table 3. Matrix design, results, and ANOVA analyses for Box-Behnken experimental design for optimization of Rep-Cap split-based rAAV production in HEK293 cell line by quadruple transfection.

Independent Variables	Coding Levels		
	-1	0	1
pHelper (c/mL)	$5.00x10^9$	$2.75x10^{10}$	$5.00x10^{10}$
pGFP_ITR (c/mL)	$5.00x10^9$	$2.75x10^{10}$	$5.00x10^{10}$
pCMV_CAP1 (c/mL)	$5.00x10^9$	$2.75x10^{10}$	$5.00x10^{10}$
pRepwt (c/mL)	$5.00x10^9$	$2.75x10^{10}$	$5.00x10^{10}$

	Variables				Response
Experimental Run	pHelper (c/mL)	pGFP_ITR (c/mL)	pCMV_CAP1 (c/mL)	pRepwt (c/mL)	rAAV1 titer (vg/cell)
1	0	0	1	-1	460.45
2	0	0	0	0	683.76
3	1	0	0	1	355.34
4	0	0	1	1	1226.63
5	0	1	0	-1	331.42
6	-1	1	0	0	133.61
7	0	0	-1	-1	115.46
8	1	-1	0	0	693.72
9	0	1	0	1	292.08
10	0	0	0	0	615.35
11	0	0	0	0	625.78
12	1	0	-1	0	111.14
13	0	0	0	0	586.68
14	0	1	1	0	1025.55
15	0	-1	1	0	1687.39
16	1	1	0	0	329.52
17	0	-1	0	-1	317.12
18	0	1	-1	0	39.86
19	0	0	-1	1	80.01
20	0	0	0	0	423.00
21	0	-1	-1	0	72.62
22	-1	0	1	0	524.50
23	1	0	0	-1	294.44
24	-1	0	-1	0	12.36
25	-1	0	0	1	187.11
26	-1	0	0	-1	39.97
27	1	0	1	0	875.61
28	-1	-1	0	0	116.02
29	0	-1	0	1	737.06

Model	Multiple R^2	*p* value [(a)]	Lack of fit test. *p* value [(b)]
rAAV1 titer (vg/cell)	0.9672	< 0.0001	0.1353
Parameters	**Coeficient**	***f Value***	***p* value [(b)]**
Constant	2.7600		
[pHelper]	0.2909	61.14	< 0.0001
[pGFP_ITR]	-0.0935	6.31	0.0249
[pCMV_CAP1]	0.5906	251.96	< 0.0001
[pRepwt]	0.1108	8.87	0.0100
[pHelper] x [pGFP_ITR]	-0.0962	2.23	0.1579
[pHelper] x [pCMV_CAP1]	-0.1829	8.05	0.0132
[pHelper] x [pRepwt]	-14.72	5.22	0.0385
[pGFP_ITR] x [pCMV_CAP1]	0.0111	0.02	0.8661
[pGFP_ITR] x [pRepwt]	-0.1053	2.67	0.1246
[pRepwt] x [pCMV_CAP1]	0.1462	5.15	0.0397
$[pHelper]^2$	-0.3518	48.32	< 0.0001
$[pGFP_ITR]^2$	-0.0567	1.25	0.2818
$[pCMV_CAP1]^2$	-0.2300	20.65	0.0005
$[pRepwt]^2$	-0.1364	7.27	0.0174

[(a)] *p values* under 0.05 are considered statistically significant with 95 % confidence, and under 0.1, statistically significant with 90 %.

[(b)] *p values* associated to lack of fit test above 0.05 mean that the hypothesis arguing that the model is suitable cannot be rejected.

The second-order polynomial model was employed to determine the optimal plasmid concentrations for rAAV1 production (**Figure 2**). Three-dimensional graphs were plotted based on the model, visually demonstrating maximal responses under varying factor conditions (**Figure 2B-G**). The analysis indicates that the pHelper and pCMV_CAP1 plasmids substantially increase rAAV1 titer, as evidenced by their high positive coefficients and highly significant *p*-values (<0.0001) (**Table 3**). In contrast, pGFP_ITR and pRep_wt, also contribute to increase rAAV1 titer but to a lesser extent. The bell-shaped response curve for the pHelper plasmid combined with pCMV_CAP1, pRep_wt and pGFP_ITR indicates an optimal concentration range that maximizes rAAV yield (**Figure 2B, C, D**). This optimal range was particularly crucial for the interaction between pHelper and pRep_wt plasmids, given the highly negative coefficient (-14.72) and significant *p*-value (0.0385) (**Table 3**). This interaction underscores the complex interplay between adenoviral helper genes and wt Rep elements when split from capsid gene expression. This model also suggests a synergistic effect between *rep* and *cap* genes, proposing that pCMV_CAP1 plasmid copy number limits rAAV yield, while Rep concentration enhances rAAV titer when capsid gene expression is already high (**Figure 2F**).

Overall, optimal plasmid concentrations for maximal rAAV1 production in HEK293 cells were: 3.86 x 10^{10} c/mL of pHelper, 1.71 x 10^{10} c/mL of pGFP_ITR, 4.84 x 10^{10} c/mL of pCMV_CAP1 and 4.31 x 10^{10} c/mL of pRep_wt. Under these conditions, the predicted rAAV1 production was 1891.9 vg/cell. Further validation experiments were performed using the optimal values determined for quadruple transfection, achieving 1636.01 vg/cell ± 156.95. In parallel, standard triple transfection was conducted achieving 4643.57 vg/cell ± 407.42. Hence, despite the superior production levels reached with standard triple transfection, the rAAV production utilizing the newly developed Rep-Cap split packaging system proved effective in HEK293 cells, generating titers that are within the same order of magnitude than those achieved by standard triple transfection. This underscores the feasibility of *rep* and *cap* gene expression in separated ORFs.

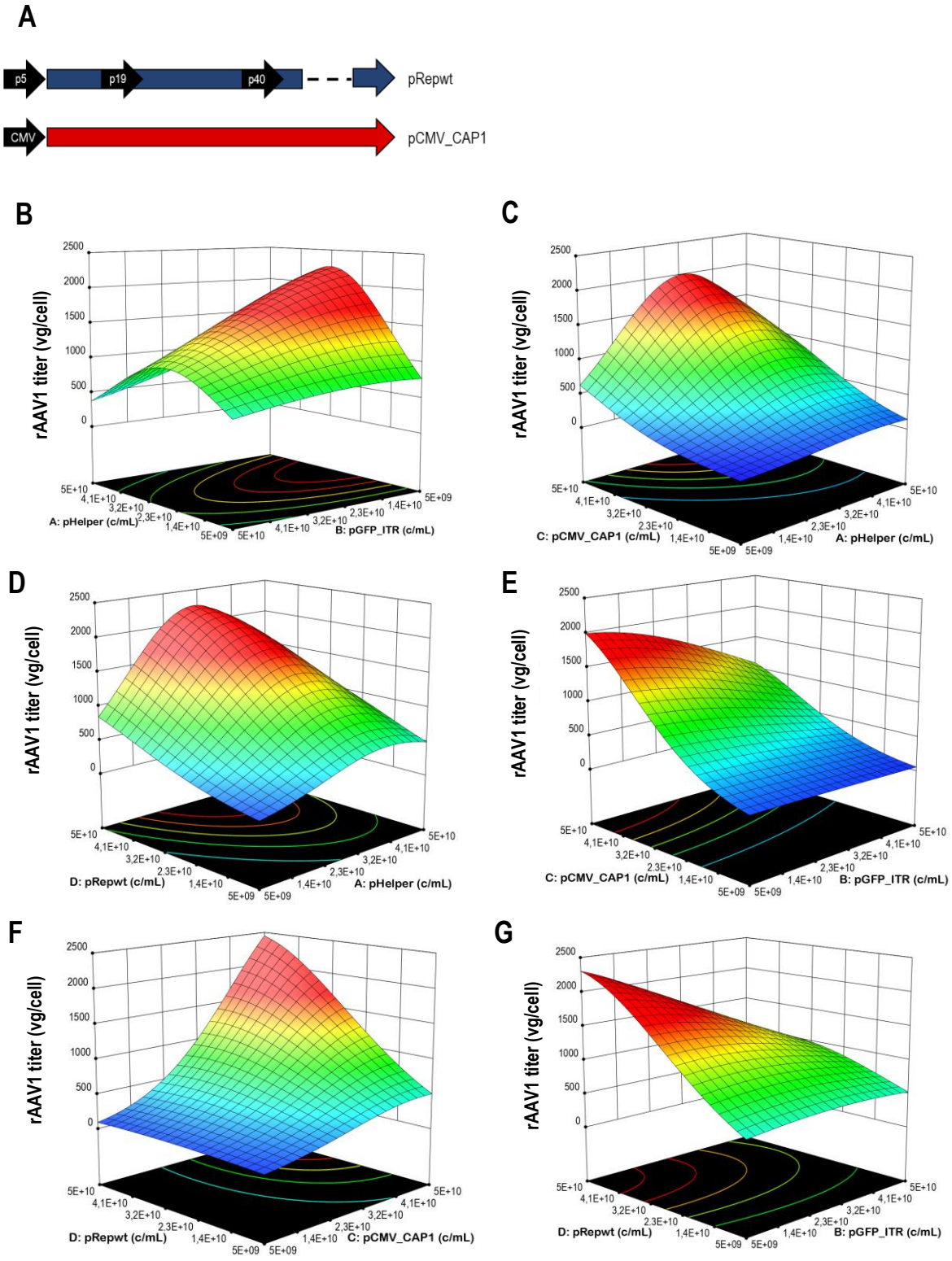


Figure 2. Rep-Cap split packaging system evaluation. A. pRep_wt and pCMV_CAP1 plasmids for Rep-Cap split packaging design for independent ORF expression. **B-G.** Response surface graphs based on Box-Behnken design of experiments results for Rep-Cap split packaging system. The Y-axis accounts for the analyzed response: rAAV1 titer (vg/cell) as a function of two independent variables presented at X- and Z-axis. The graphs were constructed by depicting two variables at a time and maintaining the other independent variables at the maximal predicted level by the model of: 3.86 x 10^{10} c/mL of pHelper, 1.71 x 10^{10} c/mL of pGFP_ITR, 4.84 x 10^{10} c/mL of pCMV_CAP1 and 4.31 x 10^{10} c/mL of pRep_wt.

Rational design of Rep inducible constructs for rAAV production

Having assessed the feasibility of the Rep-Cap split packaging system, the minimal combination of Rep proteins to drive rAAV production in HEK293 under the Tet-On inducible expression was explored. To this end, different constructs were designed incorporating rationally designed mutations (detailed in **Table 2**), aiming to individually express the four Rep proteins Rep78, Rep68, Rep52 and Rep40 as previously shown by Jalšić *et al*.,[247] and schematically shown in **Figure 3**.

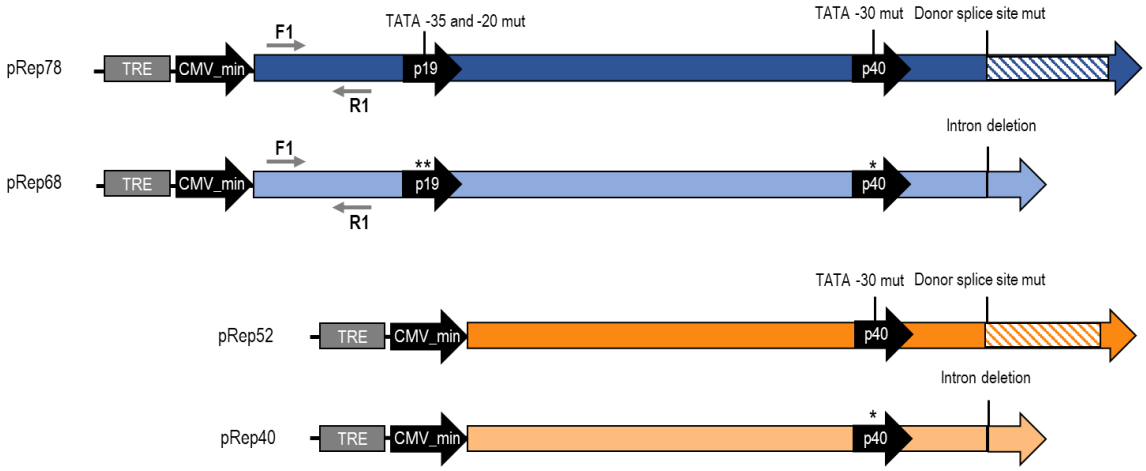


Figure 3. Schematic representation of the cassettes of the Rep inducible plasmids designed in this work. Endogenous promoter p5 (for pRep78 and pRep68) or p19 (for pRep52 and pRep40) are replaced by the tetracycline responsive element (TRE_tight_BI) and minimal CMV promoter (CMVmin), used in the Tet-On system. Primer pair 1 for RT-qPCR is depicted at its target zone. Intron is indicated in dashed lines. Symbol "*", indicates mutations compared to wt rep element.

For the constructs expressing large Rep protein, the internal p19 promoter was double mutated at -35 and -20 positions of its TATA box to avoid expression of small Rep proteins. For Rep68 expression, the intron sequence was removed. For Rep78, a mutation was introduced at the donor splicing site in order to avoid the intron splicing, a concept already described by Howe *et al.* and Jalšić *et al.*[247,385]. For the small Rep protein constructs, Rep52 and Rep40 ORFs were designed, maintaining, or eliminating the intron respectively, in line with the approach of Rep78 and Rep68. Additionally, in all Rep ORF the dispensable promoter p40 was inactivated, as it promotes capsid transcription by mutating its TATAA box motif preventing expression of non-functional or truncated viral gene products as reported by Emmerling *et al.*[204]. These mutations were designed to achieve an individual protein expression in each construct without affecting the amino acid sequence of Rep proteins.

For inducible expression of the Rep ORFs, the Advanced Tet-On® system was selected. This system is based on a mutated reverse transactivador, rtTA, which, when bound to the tetracycline-

responsive element (TRE) activates the transcription of the downstream gene in the presence of tetracycline, or its derivative doxycycline[391]. The TRE element is composed by seven tet-operator sequences (TetO) positioned upstream of the minimal CMV promoter, a shortened version of the CMV promoter (CMVmin). To note, TetO are the binding sites for the tetracycline-controlled transactivador (rtTA). Several rtTA variants have been developed to enhance system performance, reducing basal expression levels while improving inducibility[387,391,392]. Among them, rtTA2^{S}-M2 was selected, as tight control of gene expression had been previously demonstrated[388].

Here, a dual-plasmid Tet-On® system was implemented. First, prtTA was designed and constructed, containing rtTA^{2}S-M2 alongside the WPRE sequence, known to enhance mRNA stability[393]. In addition, the four designed Rep ORFs, devoid of endogenous p5 and p19 promoters were fused downstream of the CMV minimal core promoter and subsequently cloned into pTRE_tight_BI plasmid. The TRE_tight_BI was selected for its modified TRE, reported to yield high induction levels and low basal promoter activity compared with TRE, in addition, it contains a bidirectional tetracycline-responsive promoter[394]. Noteworthy, the large Rep transcripts were cloned in forward orientation of TRE, while the small Rep transcripts were inversely orientated. Envisioning a future packaging cell line, this should facilitate stable cell line generation procedure as bidirectional TRE promoter would drive both Rep large and small transcript expression simultaneously.

Functional evaluation of newly designed mutated Rep ORFs under Tet-On system

The functionality of the rationally designed large and small Rep ORFs expressed under the Tet-On® inducible system was evaluated. To this end, the four tested plasmids each expressing different Rep ORFs, were co-transfected with prtTA plasmid using equimolar quantities. The pRep_wt plasmid from the previously optimized Rep-Cap split packaging system was used as control. Doxycycline was used to activate the Tet-On® system, with parallel samples maintained without doxycycline, denoted as (+) and (-) conditions, respectively, in **Figure 4A-D**.

Initially, the inducible capacity of the different Rep constructs was evaluated in HEK293 cell line. As illustrated in **Figure 4A**, Rep78, Rep68 and Rep52 demonstrated inducible expression upon doxycycline addition, each showing higher basal expression levels compared to Rep_wt. Contrarily, Rep40 expression was remarkably lower.

Subsequently, the dependency of Rep expression on doxycycline versus E1A expression was investigated using the A549 cell line, which lacks the E1A gene. In A549 cells, inducible expression

of Rep transcripts was observed, reaching similar inducible levels to those in HEK293 cells (**Figures 4A, C**). In A549 cells higher basal levels compared to pRep_wt were detected in all constructs. This suggest that the basal expression level of the generated plasmids is not related to E1A expression within HEK293 cells, but leaky transcription from the TRE and CMVmin element. Nonetheless, the generated rep constructs were inducibly expressed upon doxycycline addition in both A549 and HEK293 cells, confirming the functionality of the generated plasmids.

Finally, the impact of the four Rep expression constructs on cell viability and viable cell density (VCD) was evaluated across both cell lines (**Figure 4B, D**). Following the addition of doxycycline, a decline in viability and VCD was observed in HEK293 cells, especially with Rep78, Rep68 and Rep52 expression. In contrast, in the A549 cell line, a less pronounced effect was noted. This differential response might be attributed to either the transfection process itself or intrinsic cellular metabolic effects (**Figure 4A, C**).

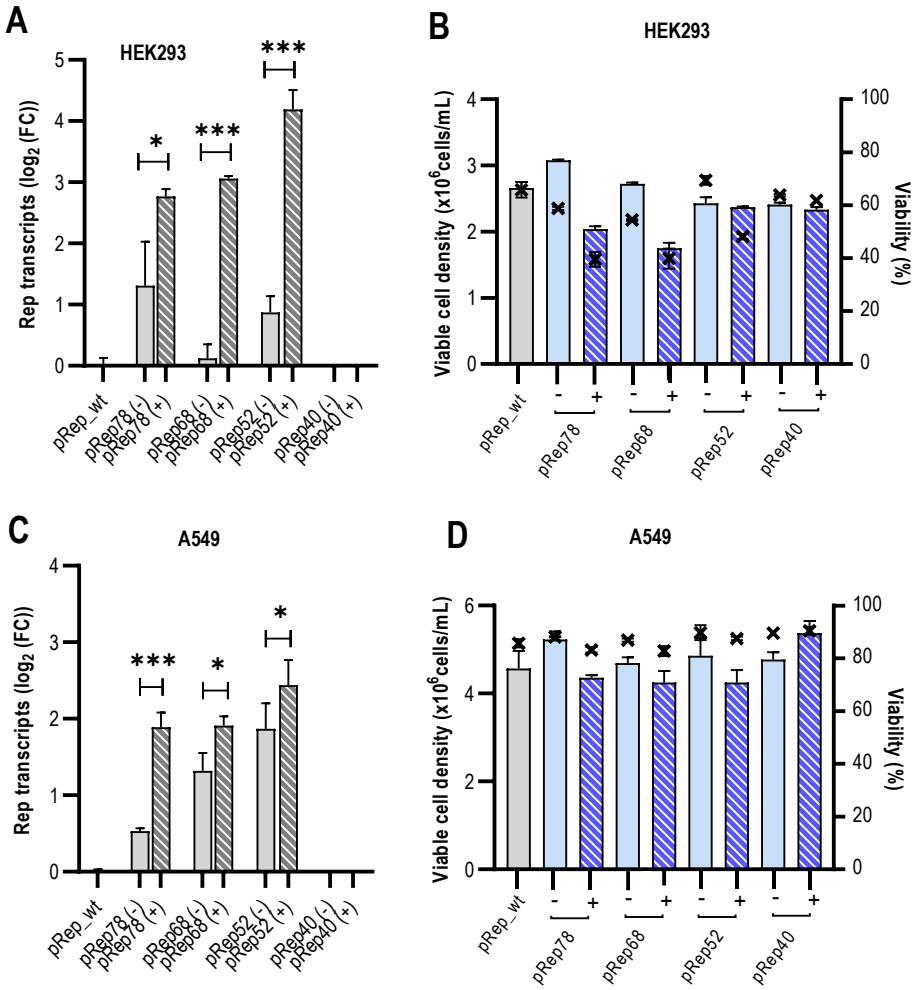


Figure 4. Functional evaluation of rationally designed inducible Rep ORFs using Tet-On system expression. Relative transcript levels of total rep using 18S as housekeeping gene measured in log_2FC using Rep_wt levels as control in HEK293 and A549 cell line respectively (**A** and **C**). Viable cell density and viability of cells at harvest in HEK293 and A549 cell line respectively (**B** and **D**). Doxycicline was added (+) or non-added (-) at 1 µg/mL. Biological triplicates were measured in triplicates. Data was analyzed by Ordinary ONE way ANOVA *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD.

Establishment of a proportional dose-response relationship between doxycycline addition and rAAV production

After confirming that the newly designed Rep ORFs were inducible following doxycycline addition, the study aimed to evaluate the optimal initial Rep plasmid concentrations at which doxycycline addition would lead to an induction in rAAV1 production. Given that Rep78 expression, driven by the pRep78 plasmid has been identified as the most toxic of all Rep proteins[57], it was selected to analyze rAAV production and monitor viable cell density at harvest. For this analysis, pRep78 was co-transfected into HEK293 cells alongside the plasmids required for rAAV production under the Tet-On system: prtTA, pHelper, pGFP_ITR, and pCMV_CAP1. Rep copy number was varied from 3.2 x 10^{10} c/mL to 3.2 x 10^{7} copies/mL (c/mL) while maintaining the concentration of the co-transfected plasmids at 3.2 x 10^{10} c/mL. Total DNA concentration was kept constant at 1 µg/mL with pMock as required. Additionally, doxycycline concentrations were varied from 0, 0.1 and 0.9 µg/mL and rAAV production was evaluated (**Figure 5**).

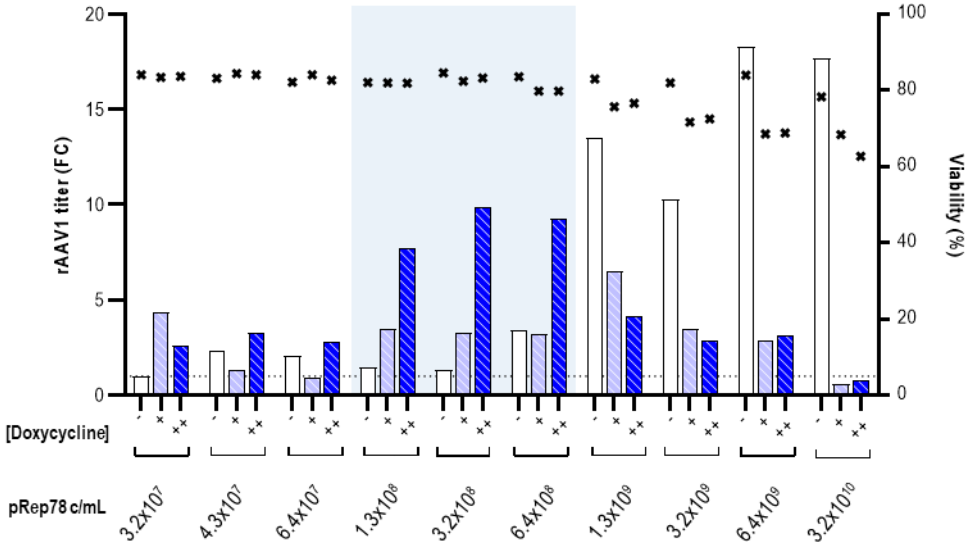


Figure 5. Dose-response relationship of doxycycline-inducible pRep78 expression on rAAV1 titer and cell viability. Impact of varying pRep78 plasmid copy numbers and doxycycline concentrations in HEK293 cells on rAAV1 titer and viability. rAAV titer is expressed as the fold change (FC) relative to the baseline control established by the lowest pRep68plus initial concentration without doxycycline addition. Different bar colors correspond to varying doxycycline concentrations (0, 0.1, and 0.9 µg/mL coded as "-", "+" and "++" respectively). The initial pRep78 plasmid concentration ranged from 3.2 x 10^{7} to 3.2 x 10^{10} c/mL. Highest plasmid concentration of pRep78 is equimolar to the other transfected plasmids: pHelper, prtTA, pCMV_CAP1 and pGFP_ITR. Proportional dose-response relationship is observed in the blue highlighted zone, where doxycycline addition leads to rAAV production.

Interestingly, a proportional relationship between the doxycycline dose and rAAV titer was observed in the range of 1.3 x 10^{8} and 6.4 x 10^{8} c/mL of pRep78 concentration (**Figure 5**). In contrast, at higher initial concentrations of pRep78, higher rAAV1 titer was observed under non-induced conditions compared to the titer achieved within the proportional dose-response range. However,

addition of doxycycline at higher pRep78 initial concentrations led to a reduction in rAAV production, showing that Rep expression levels impact on rAAV titer. Further, high Rep expression levels were associated with cellular toxicity, evidenced by a drop in cell viability following higher initial Rep concentrations upon induction (**Figure 5**).

Remarkably, the obtained results show that higher Rep expression levels do not consistently correlate with enhanced rAAV production; instead, optimal conditions must be determined. Using the developed rep-inducible system, rAAV production increased tenfold upon the addition of doxycycline. This demonstrates the effectiveness of the developed inducible Rep system in HEK293 cells for rAAV production, establishing a baseline for further system characterization.

Identification of the optimal combination of Rep constructs for maximal rAAV production under Tet-On inducible system

Having demonstrated that the newly designed pRep78 plasmid could drive inducible and proportional dose-response between doxycycline addition and rAAV production in HEK293 cell line, the next step aimed at investigating the optimal combination of the different designed Rep constructs to achieve maximal rAAV production. To this end, a DoE based on a mixed factorial design was performed, incorporating four variables at different levels. These variables were selected based on their impact on the performance of the inducible system, as identified in the previous experiments.

Variables included the total concentration of Rep plasmids at 'low' and 'high' levels, the latter corresponded to the maximum concentration of Rep78 where a proportional dose-response to doxycycline addition was observed (6.4×10^8 c/mL). Doxycycline concentration varying from 0.1 to 0.9 µg/mL was included as a variable affecting rAAV production. Moreover, to ascertain which combination of large and small Rep proteins yielded the best results, both were included as independent variables. For the large Rep proteins, Rep78 and Rep68, driven by pRep78 and pRep68 respectively, were considered. Similarly, for the small Rep proteins, Rep52 and Rep40, driven by pRep52 and pRep40 respectively, were included. Significantly, an additional '0' level indicating the absence of the respective Rep protein was introduced to determine the significance of including each Rep protein on rAAV titer.

The experimental design matrix in coded values is shown in **Table 4**. In the experimental setup, the total Rep plasmid DNA concentration varied according to the model, while the total transfected

DNA concentration was maintained at 1 µg/mL with pMock, ensuring a constant total DNA concentration in the transfection mixture.

Table 4. Matrix design, results, and ANOVA analyses for mixed-factorial design to determine optimal Rep ORF combination.

Independent Variables	Coding Levels		
	Level low	Level medium	Level high
A. Doxycycline concentration	+		++
B. Rep total quantity	Low		High
C. Rep large transcripts	0	pRep68plus	pRep68min
D. Rep small transcripts	0	pRep40plus	pRep40min

	Variables				Response
Experimental Run	A. Doxycycline concentration	B. Rep total quantity	C. Rep large transcripts	D. Rep small transcripts	rAAV1 titer (vg/mL)
1	++	low	68min	40min	$1.77x10^7$
2	+	low	68plus	0	$2.08x10^7$
3	+	low	68plus	40plus	$1.18x10^7$
4	+	low	0	0	NA
5	+	low	68min	0	$1.57x10^7$
6	+	high	0	40min	$5.43x10^5$
7	+	high	0	40plus	$4.81x10^5$
8	+	high	68plus	40plus	$1.38x10^7$
9	++	low	68min	40plus	$1.62x10^7$
10	+	high	68min	40min	$1.46x10^7$
11	++	low	0	40min	$7.84x10^5$
12	++	high	0	40plus	$3.42x10^5$
13	+	high	0	0	NA
14	++	low	0	40plus	$3.63x10^5$
15	+	high	68min	40plus	$1.40x10^7$
16	++	high	68min	0	$2.14x10^7$
17	++	high	68min	40min	$1.65x10^7$
18	++	low	68plus	0	$4.01x10^7$
19	++	high	68min	40plus	$3.40x10^7$
20	++	high	68plus	40plus	$1.59x10^7$
21	+	high	68min	0	$1.82x10^7$
22	+	low	68min	40plus	$2.02x10^7$
23	++	high	0	0	NA
24	+	high	68plus	0	$4.56x10^7$
25	+	low	0	40min	$2.62x10^6$
26	++	low	0	0	NA
27	++	low	68plus	40plus	$1.17x10^7$
28	+	low	68plus	40min	$3.50x10^7$
29	++	high	68plus	40min	$1.06x10^7$
30	++	low	68plus	40min	$2.68x10^7$
31	++	low	68min	0	$2.09x10^7$
32	++	high	0	40min	$4.60x10^5$
33	+	high	68plus	40min	$5.72x10^5$

Model	Multiple R^2	*p* value [(a)]
rAAV1 titer	0.9543	< 0,0001

Parameters	Sum of Squares [(b)]	*f Value*	*p* value
Constant	12.43		
A. Doxycycline concentration	0.0205	0.585	0.4547
B. Rep total quantity	0.0532	1.520	0.2346
C. Rep large transcripts	9.9200	141.660	< 0.0001
D. Rep small transcripts	0.1837	2.620	0.1017
B x C	0.0374	0.534	0.5956
B x D	0.1848	2.640	0.1004

[(a)] *p values* under 0.05 are considered statistically significant with 95 % confidence, and under 0.1, statistically significant with 90 %.

[(b)] Sum of squares was calculated using Type II Classifcal.

The results and statistical analysis of the mixed-factorial design are presented in **Table 4** with coefficients in terms of coded factors shown in **Table 8**. Regression analysis confirmed that the model fitted the data with a coefficient R^2 of 0.9543, indicating consistency with 95% of variability in the data. ANOVA analysis confirmed the model's statistical significance, with a Fisher's F-test p-value of <0.0001.

Table 8. Coefficients in Terms of Coded Factors of mixed factorial design in Table 4.

Parameters	Coded Factor	Coefficient Estimate [(a)]
Intercept		7.80
A. Inductor concentration	++	-0.0268
B. Rep total quantity	high	-0.0544
C. Rep large transcripts [1]	0	0.9526
C. Rep large transcripts [2]	Rep 78	0.4787
D. Rep small transcripts [1]	0	0.0750
D. Rep small transcripts [2]	Rep52	-0.0994
B x C [1]	low / 0	-0.0829
B x C [2]	low / Rep78	0.0339
B x D [1]	low / 0	0.0592
B x D [2]	low / Rep52	0.0503

[(a)] Coefficients in Terms of Coded Factors were calculated using Sum Contrasts. The coefficient estimate represents the expected change in factor value when all remaining factors are held constant.

The mixed factorial design analysis showed that only Rep large transcripts were significant to the model, with a significant *p*-value of <0.0001 (see **Table 4**). The negative coefficients of docycycline concentration and total Rep quantity suggest that these variables have a minor impact on rAAV titer (**Table 8**). For interpretation and visualization, the combination of Rep total quantity with large and small Rep factors, is presented in three-dimensional bar graphs in **Figure 6A-D**.

When large Rep transcripts were not included, no detectable production was observed (**Figure 6C, D**). Importantly, Rep78 and Rep68 addition yielded similar titers, both at low and high rep total quantity (**Figure 6A, B**). Further, similar rAAV production was observed when large Rep proteins were combined with small Rep proteins either Rep40, Rep52 or "0", indicating that small Rep addition does not further increase rAAV production (**Figure 6D**). This indicates that small Rep proteins have no influence in rAAV production under the tested conditions.

As a result, the analysis suggests that the expression of large Rep proteins alone, Rep78 or Rep68, is sufficient to support rAAV production, with both achieving comparable titers under the tested conditions.

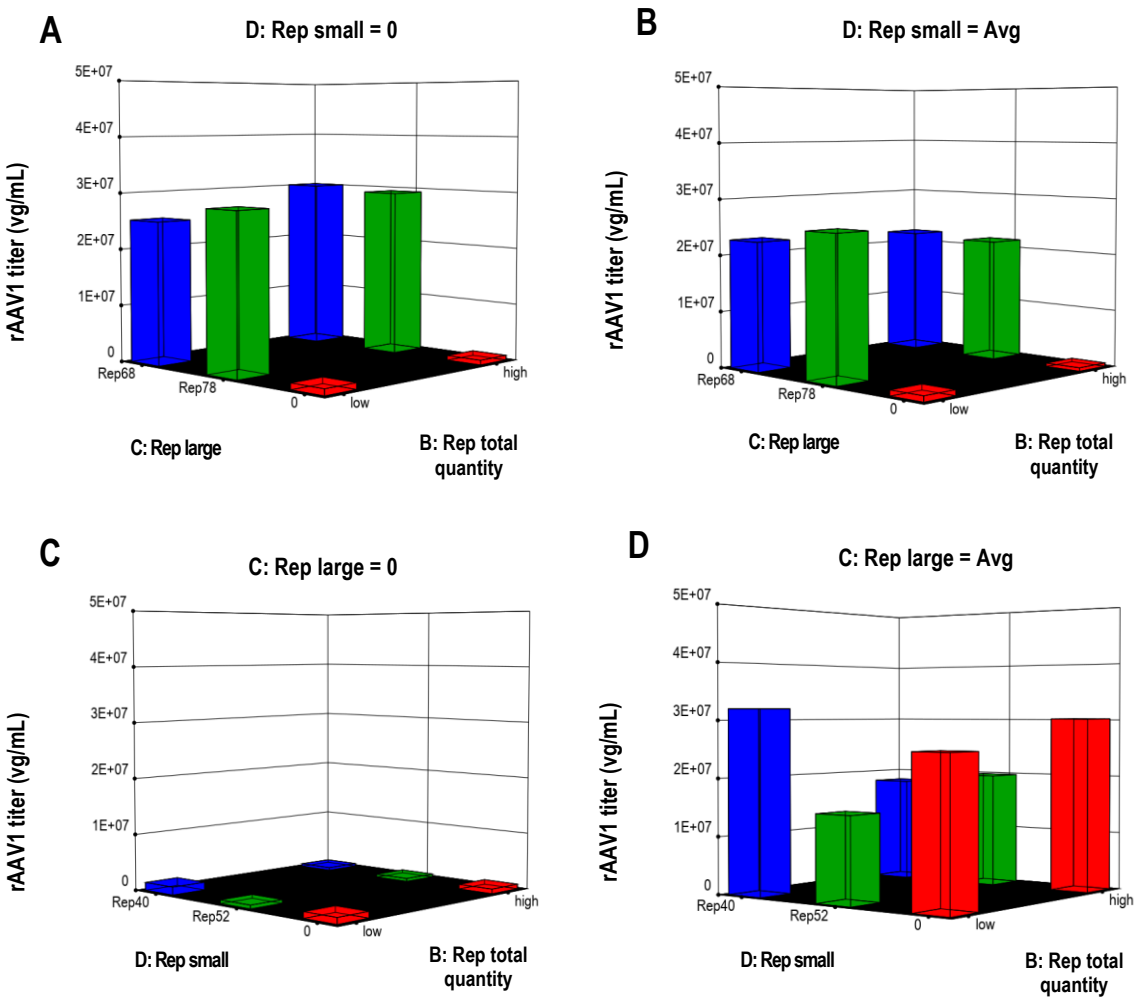


Figure 6. Results from mixed fractional Design Of Experiment to evaluate rep constructs under the Tet-On inducible system. The Y-axis accounts for the analyzed response: the rAAV titer (vg/mL) as a function of two independent variables presented at X- and Z-axis**. A, B**. rAAV1 titer in function of Rep large and Rep total quantity. Rep small variable was fixed at level "0" in A and averaged in B. **C, D.** rAAV1 titer in function of Rep small and Rep total quantity. Rep large variable was fixed at level "0" in C and averaged in D. Variable A corresponding to inductor dose was maintained at average level in all graphs.

Functional evaluation of Rep78 construct mutations

The DoE conducted under inducible conditions revealed that both pRep78 and pRep68 plasmids could support maximal rAAV production without requiring rep small-expressing constructs. These results could suggest that while mutations in large Rep constructs intended to lead to individual expression of large transcripts exclusively, it might not be sufficient to restrict completely the transcription of small Rep products. Additionally, the similar titer achieved by Rep78 or Rep68 could indicate that the ZFD contained in Rep78 intron may not play a pivotal role in the rAAV production efficiency.

To further investigate and validate these findings, a series of experiments were designed to assess the functionality of the mutations performed in the Rep78 construct. These experiments involved reverting the mutations within Rep78 individually and examining their effects on total expression levels and splicing dynamics of each mutated construct (**Figure 7A-E**).

Accordingly, three additional plasmids were constructed based on pRep78: pRep78mutS, in which the mutation blocking the splicing site was reverted; pRep78mutT, in which the two mutations within the p19 promoter's TATA box were reverted; and pRep78mutTS, which combined the reversions of both previously mentioned mutations. Reversing the mutation at the splicing site allows the splicing to occur more effectively. Further, reversing the p19 mutations decreases the control over the production of small Rep products, potentially allowing the transcription of Rep52 and Rep40 (subject to successful splicing) under the native control of the p19 promoter, as illustrated in **Table 9** and schematically represented in **Figure 7A**.

The differential quantification of each transcript variant was conducted using the primer set 1 targeting the total large transcript. Moreover, the functional evaluation of the produced species was performed using two methods: first, the “intron retention index” was evaluated with primer set 2 applying **Equation 3**. Second, the “intron splicing index” was determined using primer set 3 following **Equation 4**.

Initially, the inducibility of total large transcript expression for each construct was assessed in both the presence and absence of doxycycline. All Rep constructs demonstrated inducible expression (**Figure 7B**). Notably, pRep78mutS showed an increased expression under inducible conditions compared to pRep78. In contrast, Rep78mutT and pRep78mutTS exhibited lower basal expression levels of the large transcript compared to the pRep78 control, as depicted in **Figure 7B.**

Subsequently, the functional effects of the reverting pRep78 mutations were analyzed. Remarkable differences were observed at the non-induced conditions while at induced conditions, no significant differences among constructs were observed. The intron retention index, aimed at assessing the presence of introns within the transcripts, was then evaluated (**Figure 7C**). Under non-induced conditions, pRep78mutS was anticipated to exhibit a lower index due to native splicing efficiency (**Table 9**). However, no significant difference was observed, potentially due to incomplete splicing of pRep78mutS. In contrast, Rep78mutT, anticipated to yield both Rep78 and Rep52 proteins, showed an increased index as predicted. The outcome of pRep78mutTS was considered unpredictable due to the potential of generating all four Rep proteins.

Table 9. Summary of theoretical products achieved with the mutations reversed in each plasmid. **Figure 7A** shows schematically transcript expression expected from each plasmid. N.D: non determined.

	pRep78	pRep78_mutS	pRep78_mutT	pRep78_mutTS
Theoretical products	Rep78	Rep78, Rep68	Rep78, Rep52	Rep78, Rep68, Rep52, Rep 40
Intron retention index (2/1)	Control	↓	↑	N.D
Intron splicing index (3/1)	Control	↑	=	↑

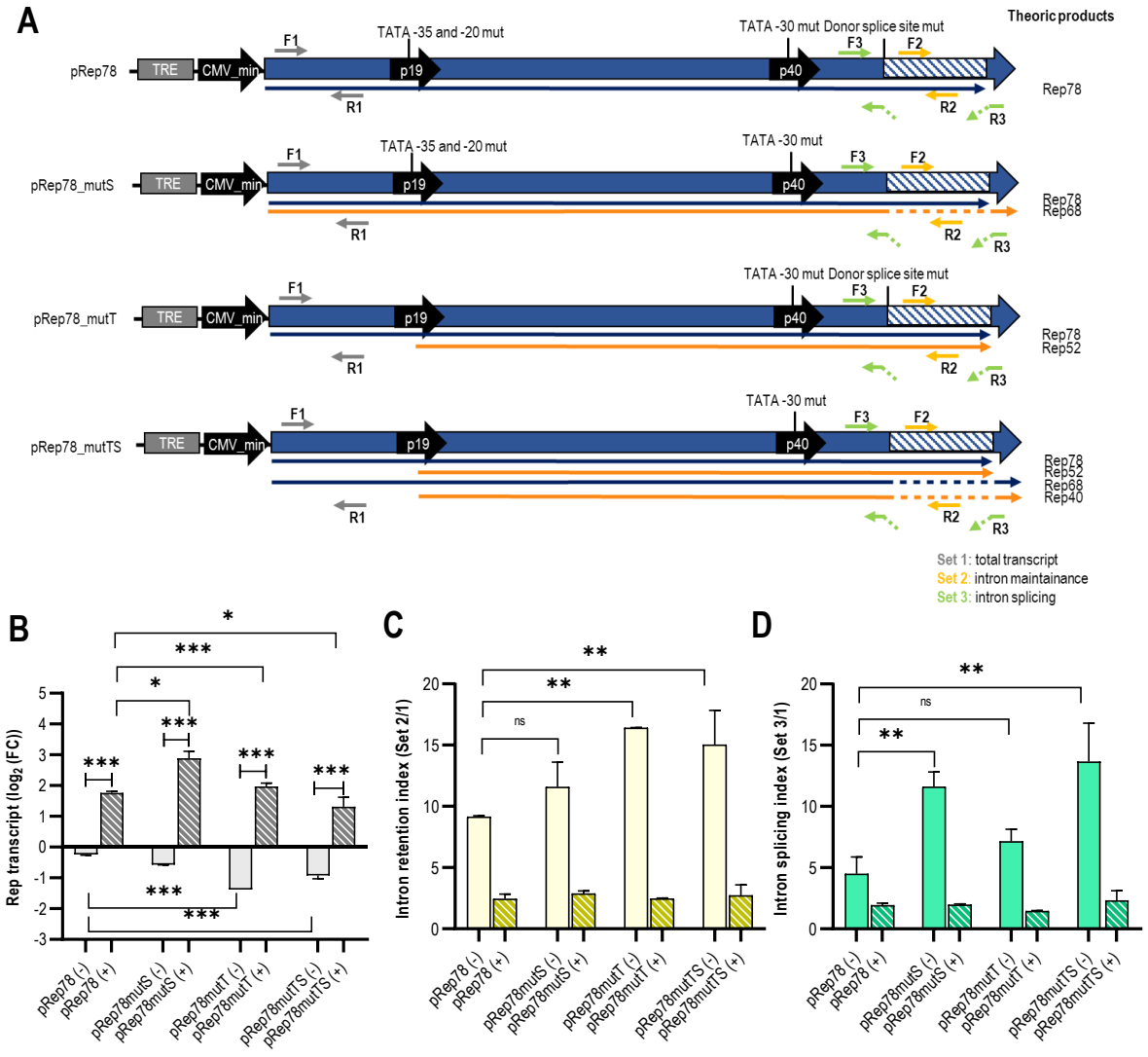


Figure 7. Functional evaluation of Rep78 mutations. A. Theoretic transcript products using the different plasmids: pRep78mutS. pRep78mutT, pRep78mutTS, generated by reverting each mutation in pRep78 control plasmid. Primer set 1, 2, 3 targeting regions are depicted. **B.** Relative total transcript levels of each evaluated construct using 18S as housekeeping gene, measured in log_2FC with pRep78(-) levels used as control **C.** Intron retention index calculated using Equation 3. **D.** Intron splicing index calculated using Equation 4. In all experiments doxycycline concentrations varied from 0, to 0.9 µg/mL coded as "-", "+" respectively).

.

To clarify the dynamics of splicing and p19 promoter leakiness, the intron splicing index was evaluated, reflecting the splicing efficiency (**Figure 7D**). At the non-induced state, pRep78mutS exhibited higher splicing index compared to pRep78 control, indicating more efficient splicing. In addition, the mutation in the p19 promoter in pRep78mutT, did not impact the splicing efficiency as expected. Finally, pRep78mutTS demonstrated an increased splicing index compared to pRep78 control at the non-induced state, indicating a higher production of spliced products. Therefore, the use of primer set 3 for splicing rate analysis proved effective in examining the functional effects of each performed mutation.

Quality and toxicity evaluation of large Rep constructs

The feasibility of generating stable cell lines with the engineered large Rep constructs was investigated by assessing rAAV production, packaging efficiency and cytotoxic effects in HEK293 cells (**Figure 8A-C**). Using a non-optimized plasmid transfection method for rAAV production, both pRep68 and pRep78 plasmids yielded similar rAAV production levels, as predicted by the performed DoE. In addition, pRep78mutS outperformed the original Rep78 in rAAV production, suggesting that endogenous splicing regulation is beneficial for rAAV production (**Figure 8B**). Moreover, the encapsidation efficiency of rAAV from all Rep mutants was quantified. Results showed consistent encapsidation levels at around 5% across all designed Rep ORFs, with pRep78mutS achieving a superior encapsidation rate of 9% under inducible conditions (**Figure 8A**).

Finally, the toxicity of the designed constructs and the effect of the introduced mutations was evaluated in terms of cell cycle progression. For this, HEK293 cells were co-transfected with the different Rep large plasmids along with prtTA at 0.5 μg/mL and analyzed relative to growth control condition. Significant alterations in the sub G1 phase were observed indicating cell cycle disruptions. In the growth control, 8.59% ± 0.04 of cells at sub G1 phase were observed. Upon Rep plasmid transfection at non-induced state, pRep78 and pRep78mutTS showed the highest toxicity and pRep68, pRep78mutT and pRep78mutS exhibited the lowest cell cycle alteration (**Figure 8C**).

Upon induction, all plasmid expression exhibited toxicity, with pRep78mutTS and pRep78 showing the most pronounced decrease on the cell cycle progression with approximately 20% of cell population at sub G1 phase (**Figure 8C**).

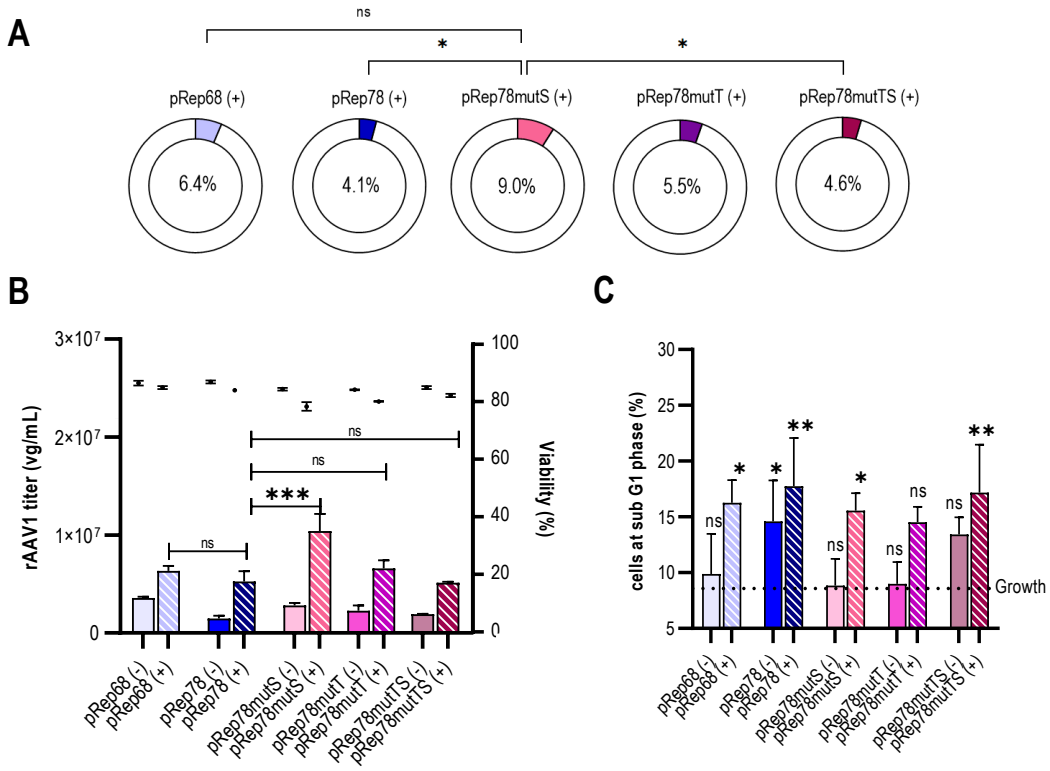


Figure 8. Quality and toxicity evaluation of pRep78 constructs. A. Full capsid % of each large Rep construct providing insights into the impact of different mutations on the encapsidation efficiency. **B.** rAAV1 production achieved using each large Rep plasmid construct, including pRep68, pRep78, and the mutated versions of pRep78: pRep78mutS, pRep78mutT and pRep78mutTS. The data is presented as rAAV titers (vg/mL) and cell viability at harvest (%). **C.** Effects of each Rep plasmid construct on cell cycle progression, specifically, sub G1 phase. In all experiments unless specified otherwise, doxycycline concentrations varied from 0, to 0.9 µg/mL coded as "-", "+" respectively).

DISCUSSION

Rep proteins play a pivotal role in various stages of viral production, including the regulation of capsid proteins, genome replication, and encapsidation[48]. Moreover, as non-structural components of the final rAAV vector, they are ideal candidates for re-design and optimization in the viral production process. Here, a rational engineering approach of *rep* gene is proposed, aiming to lay the groundwork for the future development of stable HEK293 rAAV packaging cell lines.

The split Rep-Cap packaging strategy proposed here, optimized using a DoE may offer potential improvements, not only towards packaging cell line development, but also for transient transfection systems. Emmerling *et al.*, in 2016 demonstrated that separating *rep* and *cap* genes, not only enhances rAAV titer in transient transfection but can effectively reduce recombinant competent AAV (rcAAV)[204]. Furthermore, separating Rep and Cap ORFs allows for the individual optimization of each element. In the case of capsid proteins, increased levels compared to those obtained by the wt *cap* gene have shown benefit to rAAV production[395]. Further, the split packaging system could also facilitate *rep* gene rational engineering to enhance encapsidation efficiency, demonstrated by generating *rep* hybrids[49].

Interestingly, this study demonstrates the critical influence of *rep* gene expression on rAAV titer within a Tet-On based inducible system, emphasizing the need for detailed investigation into the conditions under which rAAV production is proportional to doxycycline addition. These findings have substantial implications for the successful generation of HEK293-based *rep/cap* packaging cell line. First, the necessity of elucidating the dose-response range to effectively evaluate an inducible system's suitability for gene construction is underscored. Second, the results provided here establish important guidance for the stable integration of the rep inducible system, revealing that both *rep* copy number and inducer levels are critical factors in rAAV yield. These aspects impact clone screening and integration strategy. Previous work by Jalšić *et al*., in 2023 employed a similar approach towards the development of a HEK293-based Rep inducible cell line. However, their screening strategy of the rep-inducible system lacked proportional dose-response effect, and rAAV titer decreased upon induction. This led them to the conclusion that a combination of large and small Rep proteins was needed for rAAV production towards the generation of a stable packaging cell line[247]. In contrast, in our study, a DoE is employed based on a mixed factorial design to determine the optimal combination of the four Rep proteins towards maximization of rAAV production. This was performed under the proportional range of dose-response between inducer addition and rAAV production established in our system. The results obtained corroborate previous

ones, confirming that the presence of only small Rep proteins is not sufficient to drive rAAV production[51,56,247]. Further, non-relevant differences are observed in rAAV yield when combinations of large and small Rep proteins are used. This indicates that similar rAAV titers can be achieved whether utilizing the Rep78 or Rep68 expressing constructs. This is remarkable, as it suggests that the choice between these two ORFs may not substantially impact the efficiency of rAAV production, providing flexibility in the design of stable cell lines. This work, supports the observation that intron maintenance in Rep78 containing the ZFD is nonessential for rAAV production, as Rep68 and Rep78 led to similar production levels when induced, in agreement with previous findings[49].

Moreover, the DoE results suggest that small Rep transcript might not be needed to achieve the highest rAAV production, contrarily to what was expected as Rep52/40 have been characterized as essential for encapsidation of single-stranded AAV genomes into preformed capsids[55]. Thus, it is hypothesized that mutations introduced at p19 and splicing site of large Rep78 construct were not sufficient to tightly control the expression of small Rep transcript expression.

Here, a technical difficulty arising from the identical sequence shared by small Rep transcript to their large transcript counterpart limited the possibility of performing an absolute quantification of total small Rep transcript expression. To this end, the suitability of the introduced mutations was evaluated by reverting them and constructing different plasmids. Different primers targeting the intron and exon regions of the large Rep gene were designed, aiming to relatively analyze the splicing capacity by evaluating its spliced and un-spliced product, thereby investigating the leakiness of Rep species with and without the introduced mutations.

Notably, higher splicing rates were observed under non-induced conditions compared to those where induction was applied. Further, it was found that the addition of doxycycline to the culture led to a lack of significant differences in splicing rates among the various Rep constructs. It it has been reported that the assembly of the mRNA splicing machinery on pre-mRNA is performed concurrent with its transcription by RNA polymerase II (RNA pol II). The elongation rate of RNA pol II has been extensively characterized as a critical factor influencing pre-mRNA splicing, where an increased elongation speed may adversely affect the splicing machinery[396–398]. Based on these findings, it is hypothesized that the transcriptional elongation rate in the induced state could be linked to a non-optimal elongation rate by RNA pol II, negatively impacting the splicing process. This generates an interesting insight to be further studied in future work considering inducible systems.

Nonetheless, under non-induced conditions mutations introduced at pRep78 functioned as expected. Additionally, with doxycycline addition, both pRep68 and pRep78 achieved similar rAAV

production, coherently demonstrating that splicing of large Rep products is nonessential for maximal rAAV production.

Towards the generation of a Rep construct suitable for stable packaging cell line generation, feasibility of clonal isolation and proliferation potential at non-induced state is an important characteristic, which depends on its toxicity potential. In our study, pRep78 and pRep78mutTS plasmids showed a detrimental effect on cell cycle progression, suggesting a potential negative effect of these constructs on the obtention of viable clones. Contrarily, pRep68 and pRep78mutS emerged as promising candidates for stable cell line development with a superior capacity of packaging efficiency compared to the other constructs.

In conclusion, our results pave the way for the generation of HEK293-based packaging cell lines, capable of timely controlling *rep* gene expression while maintaining capsid gene expression at high levels consistently. Our research contributes to the development of more scalable and robust production systems. These advancements should have a substantial impact on the field of gene therapy.

CONCLUSIONS

In this study, the *rep* and *cap* gene expressions are decoupled in the Rep-Cap split packaging system to enhance control over their ORFs in rAAV production. As a first step towards the generation of a stable rAAV packaging cell line containing the *rep* gene under inducible expression, four different plasmids are designed towards inducible expression of individual Rep proteins using the Tet-On system. Characterization of this system reveals that the appropriate initial concentrations of *rep* and inducer are essential to achieve a proportional dose-response range on rAAV production. Using this insight, a DoE approach was employed in which the optimal combination of large and small Rep proteins was aimed to be characterized. However, only large Rep constructs were essential for rAAV production. This study finds that while further characterization of small transcript leakiness should be performed, the splicing site mutation and p19 mutations in large Rep78 construct were effective.

Additionally, in this work the pRep68 and pRep78mutS are characterized as suitable options for stable packaging cell line generation in HEK293 cell line. This is evidenced by maximal rAAV obtention, higher encapsidation rates and lower toxicity under non-inducible conditions.

While further system characterization is necessary, these results provide valuable insights into key first steps towards the generation of a *rep*-inducible stable packaging HEK293 cell line for rAAV production.

CHAPTER FIVE

Engineering an E1A-inducible A549 cell line for stable rAAV production

INTRODUCTION

While HEK293 cells are the standard for recombinant adeno-associated virus (rAAV) production via transient triple transfection (TT), establishing a stable cell line based on this host cell presents significant challenges due to the cytotoxicity driven by *E1A*-mediated transactivation of adeno-associated virus (AAV) *rep*[58,197] and adenovirus (AdV) *helper*[257] genes. To address this hurdle, in **Chapter four**, the potential of using a HEK293 cell line with the *rep* gene inducibly modulated was investigated towards the development of a stable rAAV packaging cell line. In this approach, the optimal Rep construct was selected through a rational modification strategy of the *rep* gene, aiming to achieve inducible rAAV production by controlling Rep protein expression using the Tet-On inducible system. Although a Rep construct with inducible control of Rep protein expression with minimal toxicity was achieved being independent on E1A expression from HEK293 cell line, low efficiency of rAAV production was observed. This may be explained by the use of non-endogenous promoters, which may be detrimental to rAAV production due to a lack of optimal temporal regulation and Rep protein expression of the viral vector production system. Overall, the development of a HEK293-based stable cell line approach for rAAV production still needs further refinement.

In addition, in **Chapter three**, the potential difficulties encountered when using HEK293 for a stable rAAV cell line generation regarding the shutdown of E1A expression by adenoviral helper proteins required for rAAV production were evidenced. Therefore, the selection of the host cell for the generation of a stable packaging/producer rAAV cell line is key. The genetic background of the cell line of choice significantly influences the strategy for genetically engineering the essential elements required for rAAV production.

Here, a novel strategy is proposed towards the development of a stable, producer cell line for rAAV production. A key aspect of this strategy is the AAV dependency on AdV co-infection for its production in wild-type (wt) infections[33,399]. Similarly, for rAAV production, helper viral proteins are needed to promote and support viral replication process[99]. The master regulator of this process is the AdV E1A protein, which plays a critical role in initiating rAAV replication by *rep* expression induction[58,197]. Additionally, E1A activates early adenovirus promoters essential for rAAV production which are *E1B, the VA RNAs, E2A,* and *E4orf6*[99,101,105–108]. Moreover, it also manipulates the host cell environment by the modulation of host cell proteins to promote viral replication and interaction with the cell cycle pathway promoting the S-phase[110]. Further, E1A expression itself can also present negative effects by triggering cell death by apoptosis, both by a

p-53 dependent and p-53 independent mechanism[111]. This is presumed to be counteracted by E1B gene expression of E1B 19K and 55K protein, both with anti-apoptotic functions[112,113].

The proposed strategy in this study involves using cell lines devoid of AdV *E1* genes as hosts for the generation of stable rAAV cell lines. Following this approach, the stable introduction and exogenous modulation of E1A gene expression can potentially trigger the rAAV production process. Such control over E1A expression in a producer cell line is expected to lead not only to the overall control of the initiation and propagation of the rAAV production process, but also to mitigate the toxic effects associated with constant viral protein expression induced by E1A transactivation. Remarkably, while this strategy may be considered as a step back into the development of producer rAAV cell line, the ability to achieve tight exogenous control of E1A may potentially streamline cell line development.

Building on this concept, the study proposes the use of A549 cell line, which was established in 1972 by D. J. Giard *et al.* from human alveolar adenocarcinomic tissue and is devoid of the E1 gene[400]. Typically, A549 cells have been cultured in adherence using animal-derived serum. However, for viral production, suspension cultures in serum-free media are preferred to facilitate process scalability. Various approaches have been explored to adapt A549 cells to suspension culture in serum-free media[401]. Back to 1994, Siegfried *et al*. adapted the A549 cell line to grow in these conditions[402], and since then several methods have been reported and patented for inducing the A549 cell line into suspension serum-free culture, underscoring its industrial interest, particularly in the production of viral particles and adenoviral vector production[401,403].

Here, an A549 cell clone derived from a suspension serum-free adapted cell bank GMP grade is used, developed by Theriva™ Biologics and currently used for adenoviral vector production. Using this cell line, the study aims to investigate the inducibility potential of E1A expression using the Tet-On system in A549 cell line in a transient gene expression approach. In this context, the capacity of A549-based cell line to produce rAAVs of high quality is evaluated and characterized.

MATERIALS AND METHODS

Mammalian cell lines and culture conditions

A serum-free suspension-adapted A549 cell line was developed and kindly provided by Theriva™ Biologics (Parets del Vallès, Spain). This cell line was derived from a current good manufacturing practice (cGMP) master cell bank available for manufacturing clinical material. HEK293 cell line (HEK293SF-3F6) is a suspension-adapted serum-free cell line developed by the National Research Council of Canada (NRC, Montreal, Canada) and kindly provided by Prof. Amine Kamen.

The medium used for suspension A549 cell culture was ExCell® CD HEK293 Viral Vector media (Merck, Darmstadt, Germany) supplemented with 6 mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA), 1x N-2 MAX Media Supplement (Bio-Techne, Minneapolis, MN, USA) and 0.6% Kolliphor (Gibco, ThermoFisher Scientific, San Jose, CA, USA).

For HEK293 cell culture, media was HyCell™ TransFx-H media (GE Healthcare, Chicago, IL, USA) supplemented with 4mM GlutaMAX™ (Gibco, ThermoFisher Scientific, San Jose, CA, USA) and 0.1% Pluronic™ F-68 non-ionic surfactant (Gibco, ThermoFisher Scientific, San Jose, CA, USA).

Suspension cell cultures were maintained in exponential growth phase in disposable polycarbonate 125mL baffled flasks with vent cap for A549 cell line or non-baffled flasks for HEK293 cells (Corning®, Tewksbury, NY, USA) at 37ºC, 5% of CO_2 and 85% RH at 130 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland). Cell concentration and viability were determined using the NucleoCounter® NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacturer's instructions. Cultures were observed under the Olympus CKX41 optical inverted microscope (Olympus, Tokyo, Japan) to qualitatively analyze cell growth and viability.

Plasmid design and generation

pGFP plasmid, which encodes for the Green Fluorescent Protein (GFP), was used as a reporter gene to monitor transfection efficiency in the cells. It was constructed by PCR amplification of pGagGFP backbone using primers containing EcoRI-HF® (R3101S, New England Biolabs (NEB), Ipswich, MA, USA) restriction digestion site as shown in **Table 1** to cut Gag polyprotein sequence and the remaining product was re-ligated using T4 DNA ligase (NEB, Ipswich, MA, USA).

pMock plasmid does not have any mammalian promoter or coding DNA sequence (CDS). It was constructed by the ligation of the pGag::eGFP backbone[350,351].

Table 1. List of primers and sequences used for each described application.

Name	Objective	Sequence (5'-3')
Fwd_GFP_EcoRI	Cloning of pGFP plasmid vector from pGagGFP	CTACAGGAATTCGATCCACCGGTCGCCAC CAT
Rev_GFP_EcoRI		GGAGGCTGGATCGGTCCGGT
Fwd_E1_EcoRI	PCR amplification of *E1* gene without E1A promoter	CGGGTCAAAGTTGGCGTTTT
Rev_E1_NotI		GCAATATCGCGGCCGCCCATAGA
Fwd_GFP_ind_mod1_NotI	PCR amplification of pGFP_ind_mod1 backbone	CCTGACGCGGCCGCCGCTCACTGACTCG CTGCGC
Rev_GFP_ind_mod1_EcoRI		TTGACAGGGAATTCGGCGACCGGTGGATC GAAGC
Fwd_GagGFP_rtTA	PCR amplification of pGagGFP backbone for clonning into rtTA gBlock	GGCCGCGACTCTAGATCATAATC
Rev_GagGFP_rtTA		TAGCGGATCTGACGGTTCACT
Rev_rtTA_gBlock_NotI	PCR amplification of rtTA gBlock	GGCTGATTATGATCTAGAGTCGC
Fwd_rtTA_gBlock		CACGCTGTTTTGACCTCCAT
Fwd_GFP_qPCR	qPCR for rAAV titer determination	AGCTGGACGGCGACGTAA
Rev_GFP_qPCR		GTGCAGATGAACTTCAGGGTCA
GFP gBlock	gBlock used for standard curve in qPCR	AGCTGTTCACCGGGGTGGTGCCCATCCTG GTCGAGCTGGACGGCGACGTAAACGGCC ACAAGTTCAGCGTGTCCGGCGAGGGCGA GGGCGATGCCACCTACGGCAAGCTGACC CTGAAGTTCATCTGCACCACCGGCAAGCT GCCCGTGCCCTGGCCCACCCTCGTGAC

Plasmids used for triple transfection (TT) were pHelper, pGG2_CMV_GFP (named pGFP_ITR hereafter) and pRep2Cap1 kindly provided by Dr. Fatima Bosch from Centre for Animal Biotechnology and Gene Therapy (CBATEG), at Universitat Autònoma de Barcelona (UAB, Bellaterra, Spain). Plasmid pRep2Cap1 encodes the Rep proteins from AAV2, and the Cap proteins from AAV1, which determine the serotype of the resulting viral vector. The plasmid pGFP_ITR, encodes for enhanced GFP under the control of the cytomegalovirus (CMV) promoter. The expression cassette is flanked by two inverted terminal repeats (ITRs) of AAV2 to enable its

packaging into AAV capsids. pHelper plasmid codifies for the adenovirus functions necessary for AAV replication and packaging, namely: *E2A, E4* and *VA RNA I* and *II*[214].

To construct the pE1_cnt plasmid, the *E1A* and *E1B* gene fragment from AdV5 (AC_000008.1) nucleotides (nt) 194 to 3509 were cloned into pcDNA3.1 + /C-(K)DYK plasmid backbone and bGh PolyA signal was inserted downstream (GenScript Europe). EcoRI enzyme restriction site was inserted into 501 position for posterior sub-cloning.

The regulation of E1 expression based on Tet-On® system with doxycycline addition consists of two mammalian expression vectors. For the construction of the plasmid coding for constitutive expression of the reverse transactivator rtTA Advanced (Clontech, TakaRa, San Jose, CA, USA), the open reading frame (ORF) containing rtTA-Advanced ($rtTA2^{S}$-M2)[388] and WPRE was obtained as a gBlock at IDT (Integrated DNA Technologies, Leuven, Belgium) flanked by regions with sequence complementarity to a backbone containing CMV promoter and SV40 polyA signal (from pGagGFP plasmid). Both the gBlock coding for rtTA and pGagGFP were PCR amplified using primers shown in **Table 1**. Products were checked by gel electrophoresis and then incubated with Hi-Fi Master Mix (E2621S, NEB, Ipswich, MA, USA) for 1 hour at 50ºC using a molar ratio of 1:2.

To construct the tetracycline (TET)-responsive plasmid for AdV5 E1A and E1B expression, the *E1A* and *E1B* gene fragments were cloned downstream of the TRE2 operator adapted from Clontech (TakaRa, San Jose, CA, USA) backbone sequence as in Bertrand Collet's plasmid #52420 from Addgene[404]. In short, pTRE2_GFP_ind_mod1 was cloned in house with a modified multiple cloning site (MCS) downstream of the TRE2 promoter and CMV minimal promoter designed to incorporate enzyme restriction sites suitable for cloning E1A and E1B expression construct. The E1A promoter region was selected according to Qiao *et al.*[257], Zheng *et al*.[363] and Emmerling *et al*.[204]. The E1A coding region without E1A promoter (nt 501 to 3509 from AC_000008.1) and backbone from pTRE2_GFP_ind_mod1 were amplified by PCR using primers shown in **Table 1**. PCR products were checked by gel electrophoresis and then both fragments were digested with EcoRI-HF® (R3101S, NEB, Ipswich, MA, USA) and NotI-HF® (R3189S, NEB, Ipswich, MA, USA). Fragments were ligated using T4 DNA ligase (M0202S, NEB, Ipswich, MA, USA). pE1_ind plasmid was obtained, in which E1A expression is driven by TRE promoter and E1B 19K and 55K products by its endogenous promoter.

Plasmids were confirmed by Sanger sequencing (Servei de Genòmica, UAB, Bellaterra, Spain) and amplified in a chemicompetent *E. coli* DH5α strain produced *in house* growing the cells in LB medium (Conda, Madrid, Spain) supplemented with 10 µg/ml kanamycin (Sigma-Aldrich, St. Louis, MO, USA) or 100 µg/ml ampicillin (Sigma-Aldrich, St. Louis, MO, USA) depending on the *E. coli*

antibiotic resistance present on each plasmid. Plasmid purification was carried out using the Endofree Plasmid MIDI kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The quantity of isolated DNA was assessed using Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, San Jose, CA, USA).

Standard transient transfection protocol

Transient transfection of A549 suspension cells was performed using 25-kDa linear polyethilenimine (PEI) as transfection reagent (PEIpro®, Polyplus-transfection SA, Illkirch-Graffenstaden, France). The standard transfection protocol was the following. Briefly, plasmid DNA was diluted in non-supplemented fresh culture media (10% of the total culture volume to be transfected) and vortexed 10 seconds. Transfection reagent PEI was added to the diluted DNA mixture, vortexed three times for 3 seconds and incubated for 15 min at room temperature. The DNA:PEI complexes were then added to the cell culture. DNA and PEI concentrations were adapted for each experimental requirement.

Transient transfection optimization using Design of Experiments (DoE)

Transient transfection on A549 cells was optimized by Design of Experiments (DoE). A Response Surface Methodology (RSM) was applied to assess the effect of the main factors influencing transient transfection: cell density at the time of transfection, DNA/cell ratio, and PEI/cell ratio. For transfection efficiency analysis, pGFP plasmid was used. The selected DoE responses were transfection efficiency, measured by the GFP expressing cells 48 hpt (% GFP positive cells) and cell viability (% of viability). The percentage of GFP positive cell response was transformed to *log* values using the logarithmic regression to reduce the funnel shape effect of the residuals. To define the influence of the studied variables on each evaluated response, a Box-Behnken design was selected. The three variables were screened at three levels: a low level coded as -1, a medium level coded as 0 and a high level coded as +1, as indicated in **Table 2**. Box-Behnken experimental results were fitted to a second-order polynomial equation by the type III sum of squares method (**Equation 1**):

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i \cdot X_i + \sum_{i=1}^{k} \beta_{ii} \cdot X_i^2 + \sum_{i=1}^{k} \sum_{j>1}^{k} \beta_{ij} \cdot X_i \cdot X_j \qquad (1)$$

where Y is the response variable under consideration (GFP-positive cells (%) or viability (%) at harvest), β_0 is the model intercept term, β_i is the linear coefficient, β_{ii} is the quadratic coefficient,

β_{ij} is the interaction coefficient, and the coded factors X_i and X_j are the independent variables of study. This equation facilitates the exploration of *k* factors at three levels, and it is used to predict the optimal values of the independent variables.

Quadruple transfection for rAAV production optimization by DoE

To optimize the quadruple transfection for optimal rAAV production in the A549 cell line, a four-factor, three-level Box-Behnken experimental design was employed. The studied variables were the four essential plasmids, including: pHelper, pRep2Cap1, pGFP_ITR and pE1_cnt. Each of these variables were screened across three different levels as shown in **Table 3**. **Equation 1** was used to model the relationship between the independent variables and the response (rAAV titer). In this context, k=4, corresponding to the four independent variables considered in this study.

DoE statistical analysis

Statistical analysis and three-dimensional graphs of the different DoEs facilitating model interpretation was performed using the Design-Expert® Software (Stat-Ease, Inc, Minneapolis, MN, USA). The quality of the fit of the model equation is expressed by the coefficient R^2. The lack-of-fit (LOF) test was used to evaluate differences between experimental and pure error of the fitted equations. The overall significance of the model was determined by analysis of variance (ANOVA) *F*-test, whereas the significance of each coefficient was determined by the corresponding *t*-test. *p*-values of 0.05 and 0.1 were considered as statistically significant with 95% and 90% confidence, respectively.

Flow cytometry

The percentage of GFP positive cells was assessed using Cytoflex (Beckman Coulter Inc., Brea, CA, USA) at Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria (UAB, Bellaterra, Spain). First, SSC-H vs FSC-A and FSC-A vs FSC-H density plots were generated to gate each individual cell population. Then, GFP positive population was assessed in Y585-PE-A vs GFP B525-FITC-A density plot. Gating was manually adjusted with A549 non-transfected cells as a control. Data was processed using CytExpert v.2.3 software (Beckman Coulter Inc., Brea, CA, USA) (RRID:SCR_017217).

Transient transfection and rAAV production

Transient transfection for rAAV production in A549 cell line was performed following the standard protocol detailed earlier. For rAAV production in A549 cells by quadruple transfection, Box-Behnken was used to optimize each plasmid concentration. The optimal conditions are detailed in the **Results** section.

To test the Tet-On inducible system in transient transfection conditions, optimized parameters of cell density, DNA and PEI concentrations were used. Transfections were carried out using an equal copy number of each of the transfected plasmids and pMock was used to target the desired DNA concentration. When required, according to planned experiments, culture was supplemented with doxycycline every 24 hpt (D5207, Avanti, Sigma-Aldrich, San Luis, MO, USA) dissolved in water.

For rAAV production in HEK293, transfections were carried out at a cell density of 2 x 10^6 cells/mL. Polyplexes were formed as previously described. PEIpro® was added in 1:2 (w/w) DNA:PEI ratio, with a total DNA concentration of 1 µg/ml. Equal plasmid copy number for each of the triple transfection plasmids was used: pHelper, pRep2Cap1, pGFP_ITR and pMock.

Cell culture for rAAV1 harvest was performed at 72 hpt. Cell culture was centrifuged at 1000 x g for 10 minutes, the supernatant was collected, and cellular pellets were stored. For cell lysate collection, cells were treated with 1x cell lysis solution (500 mM Tris, 1 % Triton X-100 V/V, 20 mM $MgCl_2$ at pH 7.5) and incubated at 37ºC, 130 rpm, for 1 hour. Lysate mixture was centrifuged to separate cell debris at 4000 x g for 20 minutes and supernatant was collected as cell lysate fraction. All fractions were stored at -80ºC until analysis was performed.

Quantitative polymerase chain reaction (qPCR) for rAAV quantification

For the determination of physical vector genome titer of rAAV (vg), 5 µL of the lysate fraction from each sample were added to DNase I (0.066 mg/mL, 10104159001, Roche, Basel, Switzerland) prepared in DNase I buffer (0.16 mM $CaCl_2$, 6.6 mM $MgCl_2$, 17.3 mM Tris-HCl, 0.066 % Pluronic F68, pH 8.0). This mixture was incubated at 37 ºC for 16 hours. Then, 4 µL of 30mM EDTA (AM9260G, Invitrogen, Carlsbad, CA, USA, in 10 mM Tris-HCl, pH 8.0) was added to each tube and incubated at 70ºC for 30 minutes to inactivate DNase I. After, 5 µL of Proteinase K (2mg/mL, EO0491, ThermoFisher Scientific, San Jose, CA, USA) was added, incubated at 55ºC for 2 hours and inactivated at 95ºC for 15 minutes. The quantification of viral genomes was conducted using quantitative PCR (qPCR). This reaction was set up using iTaq Universal SYBR Green Supermix (1725124, BioRad, Hercules, CA, USA) and a set of primers specific to a region within the eGFP gene (**Table 1**). Thermocycling profile was based on manufacturer's recommendations in a

CFX384 Touch Real-Time PCR Detection System (**Table 4)**. For preparing the standard curve, a gBlock (IDT, Leuven, Belgium) containing the targeted eGFP region of known concentration was used (sequence shown in **Table 1**). A serial dilution of this standard, ranging from 10^8 to 10^2 copies/µL) was prepared and amplified by qPCR. The concentration of AAV samples was the determined by interpolating their amplification data in the standard curve. Additionally, an internal rAAV1-CMV-GFP reference was amplified alongside the samples in all the quantifications.

Table 4. Conditions of cycling temperatures and time for qPCR and RT-qPCR**.**

Reaction	Cycles (temperature and time)
qPCR for rAAV quantification and RT-qPCR	95ºC, 3 min + 40 x [95ºC, 10 sec + 60°C, 30 sec] + 95ºC, 10 sec + 65 to 95°C, 0.5ºC/5 sec + 4°C hold

Reverse transcription quantitative PCR (RT-qPCR)

Total RNA was extracted from cell pellets using the Promega Maxwell RSC simplyRNA Tissue Kit (Cat No. AS1340) according to the manufacturer's protocol and stored at -80ºC. Quantity of isolated RNA was assessed using Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, San Jose, CA, USA). cDNA was synthesized using 1 µg of total RNA per sample with the iScript cDNA Synthesis Kit (BioRad, Hercules, CA, USA).

Reverse transcription quantitative PCR (RT-qPCR) assays were performed with iTaq Universal SYBR Green Supermix (1725124, BioRad, Hercules, CA, USA) CFX384 Touch Real-Time PCR Detection System. Each qPCR reaction mixture contained 2 µl of 5x diluted cDNA and a primer concentration of 500 nM. The qPCR cycles used for both target gene and ribosomal 18S internal control are provided in **Table 4**. At the end of the PCR, a melting curve was recorded from 65 to 95ºC. The qPCR assay was performed in triplicate for all samples. Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method[365]. Briefly, the target gene Ct values for each sample were subtracted the mean Ct value of 18S of each sample to obtain ΔCt. Then, ΔΔCt was calculated by subtracting the ΔCt of target gene to the ΔCt of reference control sample. Fold changes in relative gene expression were calculated by $\log_2(2^{-\Delta\Delta Ct})$. All conditions were performed in triplicate. Primer sequences are provided in **Table 5**.

Table 5. List of primers for RT-qPCR.

RT-qPCR primers	Sequence (5'-3')
Fwd_E1a_total	TCCGGTCCTTCTAACACACC
Rev_E1A_total	GGCGTTTACAGCTCAAGTCC
Fwd_13S_12S	TTTTGAACCACCTACCCTTC
Rev_13S	CCACAGGTCCTCATATAGCAAA
Rev_12S	GGAGTCACAGCTATCCGTACTACT
Fwd_18S	TCAAGAACGAAAGTCGGAGG
Rev_18S	GGACATCTAAGGGCATCACA

rAAV1 purification

Cell lysate was treated with Benzonase ® (Merck Millipore, Kankakee, IL, USA, 25 U/mL, 1 mM Tris, 0.20 mM $MgCl_2$, pH 8) in 5 mL of volume for 1 hour at 37ºC, 130 rpm to reduce DNA contamination. After 30 minutes of incubation, $MgS0_4$ was added at 0.0375 M to decrease rAAV aggregation. Then, 1/10 of the volume of chloroform was added and the sample was vortexed for 5 minutes. After that, NaCl was added to each tube at a final concentration of 1M and vortexed for an additional 10 seconds. Samples were centrifuged 5 minutes at 3000 x g and the aqueous phase was collected and transferred into 50 mL conical tube. To further reduce protein impurities, chloroform was added at 1:1 (v:v) to each tube, vortexed for 10 seconds and centrifuged 5 minutes at 3000 x g twice until aqueous phase was clear. Aqueous phase was collected, and tubes were left open in a biosafety cabinet to evaporate chloroform for 45 minutes. Sample was then transferred to Amicon Ultra-4 Centrifugal Filter (Merck Millipore, Kankakee, IL, USA) and tubes were centrifuged 5 minutes at 4000 x g. Then, elute was discarded, and 1 mL of PBS was added to the retentate for sample buffer exchange through centrifugation. This step was repeated twice to wash the concentrate. Finally, samples were sterile filtered at 0.22 µm and stored at -80ºC until analysis was performed. Samples were treated in triplicate and kept on ice during all the purification procedure. Centrifugations were performed at 4 ºC.

rAAV *in vitro* potency assay

HEK293 cells were seeded at 1 x 10^5 cells/well in a 96-well plate and incubated for 1 hour at 37ºC until cells adhered. Then, rAAV1 preparations were normalized to 1.2 x 10^9 vg/mL and added to each well at different dilutions with MOIs ranging from 100 to 300, in triplicate. The dilutions used were determined experimentally to yield less than 50% of GFP-positive cells, avoiding inaccuracies caused by one cell receiving multiple viral charge[352,353]. Cells were harvested 48 hours post-

transduction, and the percentage of GFP-positive cells was assessed using Cytoflex (Beckmn Coulter Inc., Brea, CA, USA) at Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria (UAB, Bellaterra, Spain). First, SSC-H vs FSC-A and FSC-A vs FSC-H density plots were generated to gate individual cell population. Then, GFP-positive population was assessed in Y585-PE-A vs GFP B525-FITC-A density plot. Gating was manually adjusted with HEK293 non-transduced cells as a control. Samples were concentrated 2-fold before analysis. 5000 events were analyzed per sample at a flow rate of 70 µL/min. Results were treated with the CytExpert v.2.3 software. The TU titer was calculated using the following formula (**Equation 2**):

$$\text{Transducing units (TU)/mL} = \frac{\text{GFP x N x FD x 1000}}{\text{V}} \qquad (\mathbf{2})$$

where GFP is the number of GFP-positive cells per well, N is the number of cells per well at the time of transduction (corresponding to 1 x 10^5 cells per well), FD is the fold dilution of vector sample used for transduction at different MOI's, V is the volume of diluted vector sample added into each well for transduction in microliters[352,353].

Fluorescence microscopy analysis

Samples were placed in 35 mm glass bottom Petri dishes with 14 mm microwell (MatTek Corporation, Ashland, MA, USA) prior to its visualization under the microscope. Imaging of cells transduced by rAAV1 particles was performed using Olympus FV1000 confocal fluorescence microscope (Olympus, Tokyo, Japan) at Servei de Microscòpia (UAB, Bellaterra, Spain).

Total capsid determination by ELISA

rAAV1 capsid titters were determined by ELISA (RAAV1 Titration ELISA, cat# PRAAV1, PROGEN Biotechnik GmbH, Heidelberg, Germany) according to the manufacturer's instructions. Samples were serially diluted to fit the standard curve of reference control provided by the kit. The standard curve was generated with the standard kit provided, using a second-order quadratic polynomial curve fit and sample concentrations were automatically interpolated from the standard curve. Absorbances were measured at 450 nm (SPECTROstar Nano, Ortenberg, Germany). Concentration values were averaged across the dilutions that were within the adjusted absorbance range recommended by the vendor and the standard deviation was calculated. A minimum of two and to up to three dilutions were used for each sample with a coefficient of variability of less than 15%[341].

Transmission electron microscopy (TEM)

rAAV1 preparations were analyzed after air-dried negative staining. Briefly, rAAV1 samples from after purification protocol were deposited onto carbon-coated copper or Holey carbon 200 mesh grids (Micro to Nano, Wateringweg, the Netherlands). Thereafter, 8 µL of sample were loaded onto the grid and incubated at room temperature for 1 min. Excess sample was carefully drained off the grid with the aid of filter paper. Samples were negatively stained with 8 µL of 2% w:v uranyl acetate by incubation at room temperature for 1 minute[367]. Excess stain was drained off as previously indicated and grids were dried at room temperature until analysis. TEM examinations were performed with a Jeol JEM-1400 (JEOL USA, Pleasanton, CA, USA) transmission electron microscope equipped with an ES1000W Erlangshen charge-coupled device camera (CCD) (Model No. 785; Gatan, Pleasanton, CA, USA) at Servei de Microscòpia (UAB, Bellaterra, Spain).

RESULTS

Optimization of A549 transient transfection using Design of Experiments

The A549 cell line used, adapted for suspension and serum-free growth, exhibited low transient transfection efficiency. This presented a challenge towards evaluating its rAAV production capacity by transient transfection.

To this end, a DoE approach was performed to study the variation of key variables in transient transfection process efficiency in A549 cells: DNA to cell ratio, PEI to cell ratio, and cell concentration at the time of transfection. To obtain statistically relevant information while minimizing the number of experimental runs required, a Response Surface Methodology (RSM) employing a Box-Behnken design was selected as an efficient methodology demonstrated previously by Gutiérrez-Granados *et al*.[405].

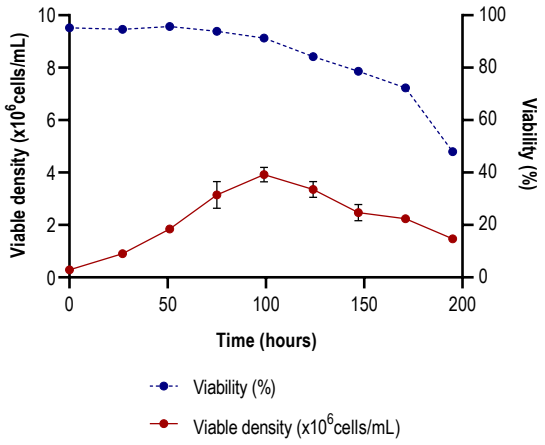


Figure 1. Growth kinetics of A549 cells. A549 growing at exponential phase were seeded at 0.3x10^6 cells/mL in batch in a 125-mL baffled flask in triplicate. Cell culture density and viability were determined every 24 hours. Error bars indicate SD of triplicates.

To determine the DoE design space regarding cell density at the time of transfection, the cell growth kinetics of A549 cell line were analyzed. The exponential growth phase was reached between 0.3 – 3.2 x 10^6 cells/mL and maximum cell density achieved was 4.02 ± 0.2 x 10^6 cells/mL after 96 hours of cell culture (**Figure 1**). The exponential growth phase is well-established as the optimal cell state for efficient transfection, thus, the cell density range for transfection was set from 1 to 3 x 10^6 cells/mL[287]. Regarding DNA to cell ratio and PEI to cell ratio, the DoE design space were set based on literature review and previous experience with this system[405]. The DNA to cell ratio was established at 0.1 to 1.9 pg/cell, and pGFP plasmid was used to facilitate transfection efficiency characterization. PEI to cell ratio was set at 0.2 to 3.8 pg/cell, with the upper limit constrained by the potential cellular toxicity of the transfection reagent[406].

Table 2. Matrix design, results and ANOVA analysis of Box-Behnken experimental design of the different variables for transient transfection optimization in A549 cells.

Independent Variables	Coding Levels		
	-1	0	1
Cell density (x10^6cells/mL)	1	2	3
DNA/cell ratio (pg/cell)	0.1	1	1.9
PEI/cell ratio (pg/cell)	0.2	2	3.8

	Variables			Responses	
Experimental Run	Cell density	DNA/cell ratio	PEI/cell ratio	GFP positive cells (%)	Viability (%)
1	0	0	0	5.68	85.95
2	1	0	-1	0.07	96.45
3	-1	0	-1	0.04	98.16
4	0	0	0	10.10	75.32
5	-1	-1	0	6.52	64.17
6	0	0	0	6.00	71.15
7	-1	1	0	18.00	75.94
8	0	0	0	6.43	79.66
9	1	1	0	18.87	82.17
10	-1	0	1	19.05	35.95
11	0	0	0	5.29	68.72
12	0	-1	-1	0.05	65.96
13	1	0	1	28.50	16.94
14	1	-1	0	18.00	54.19
15	0	1	1	35.00	45.59
16	0	1	-1	0.09	91.50
17	0	-1	1	11.80	37.18

Model	Multiple R^2	*p* value [(a)]	Lack of fit test, *p* value [(b)]
A (GFP positive cells - based)	0.9959	< 0.0001	0.6471
B (Viability - based)	0.9052	0.0075	0.0984

	Model A		Model B	
Parameters	Coeficient	*p* value [(b)]	Coeficient	*p* value [(b)]
Constant	0.8137		76.16	
[Cells] [(c)]	0.1099	0.0179	-3.06	0.4424
$[Cells]^2$	0.1486	0.0152	-2.61	0.6295
[DNA]	1.2800	0.0043	9.21	0.0441
$[DNA]^2$	-0.1051	0.0083	-4.43	0.4205
[PEI]	-0.010	< 0.0001	-27.05	0.0002
$[PEI]^2$	0.0542	< 0.0001	-11.67	0.0590
[Cells] x [DNA]	0.1573	0.0761	4.05	0.4707
[Cells] x [PEI]	0.1791	0.7462	-4.32	0.4427
[DNA] X [PEI]	-0.9255	0.3190	-4.28	0.4471

[(a)] *p values* under 0.05 are considered statistically significant with 95 % confidence, and under 0.1, statistically significant with 90 %.

[(b)] *p values* associated to lack of fit test above 0.05 mean that the hypothesis arguing that the model is suitable cannot be rejected.

[(c)] [Cells] Cell density (x10^6cells/mL)

A three-factor, three-level Box-Behnken design was performed using the selected concentration ranges for each variable to define the design space boundaries. A 17-experiment matrix was defined in which the central point was performed in quintuplicate to assess the experimental error (**Table 2**). The response variables were the percentage of GFP-positive cells, analyzed in Model A and cell viability at the time of harvest, analyzed in Model B (**Table 2**). To identify the optimal transient transfection conditions for A549 cells, variable levels leading to maximal GFP positive cells in Model A were further analyzed in the context of Model B. This approach was employed to ensure that the optimal conditions of transient transfection efficiency did not compromise cell viability.

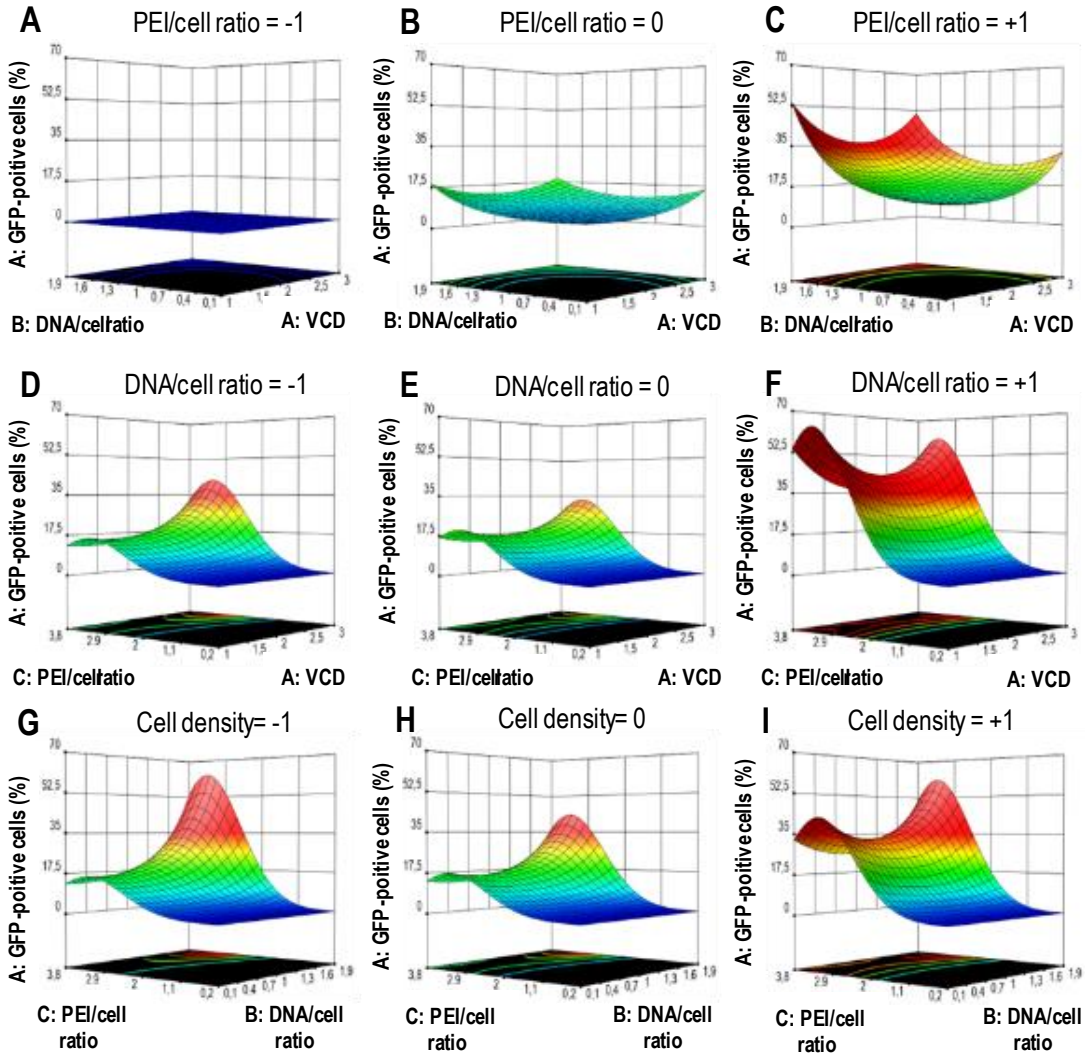


Figure 2. Response surface graphs based on Box-Behnken design of experiments results from Model A analyzing transfection efficiency. The Y-axis accounts for the analyzed response: the percentage of cells expressing GFP at 48 hours post-transfection (hpt) as a function of two independent variables presented at X- and Z-axis. **A, B, C** DNA/cell ratio vs Cell density; **D, E, F**. PEI/cell ratio vs Cell density; and **G, H, I.** DNA/cell ratio vs PEI/cell ratio. Graphs were constructed by depicting two variables at a time and maintaining the third independent variable at a fixed level (-1, 0, 1), as indicated above each individual surface response graph. VCD, Viable cell density.

The experimental design matrix, responses, and statistical analysis are presented in **Table 2**. Data were fitted to a second-order model by linear regression analysis for each response variable. For Model A (GFP-positive cells), regression analysis showed a robust fit to the data, with a coefficient R^2 of 0.9959, reflecting consistency with 99% of the data variability. The model's statistical significance was validated by ANOVA, with a Fisher's F-test producing a *p*-value of <0.0001, confirming the model's significance. The lack of fit was found to be non-significant (*p*-value 0.6471). Similarly, for Model B (cell viability), regression analysis indicated a fit with an R^2 of 0.9052, consistent with 90% of the data variability. ANOVA further confirmed the model's statistical significance with a Fisher's F-test *p*-value of 0.0075. The lack of fit for this model was also non-significant (*p*-value 0.0984).

First, the second-order polynomial model from Model A, which analyzed transient transfection efficiency, was used to generate response surface graphs and study the interactions between DNA, PEI and viable cell concentration on GFP positive cells (**Figure 2A-I**). All independent variables significantly influenced transfection efficiency, being PEI the most relevant, followed by DNA, and to a lesser extent, cell density at the time of transfection as evidenced by their respective *p*-values (**Table 2**).

In the three-dimensional graphs, maximum GFP positive cells were achieved within the cell density range of 1 to 3 x 10^6 cells/mL when variables were kept at the upper range of the design space (**Figure 2G-I**). The PEI to cell ratio showed a pronounced impact on transfection efficiency, with outcomes ranging from 0 to nearly 40% GFP positive cells (**Figure 2A-C**). Similarly, the DNA to cell ratio impacted transfection efficiency, with the higher ratio favoring transfection (**Figure 2D-F**). Overall, the optimal conditions for enhanced transfection efficiency were located near the higher design space boundaries (**Figure 2C, F, I**).

Next, the influence of the studied parameters on maximal transfection efficiency (Model A) were studied in the context of cell viability at harvest as the response in Model B (**Table 2**). To this end, the second-order polynomial model from Model A and Model B were used to predict GFP positive cells and cell viability values, respectively, by fixing two variables and examining the influence on the third. This dual-model approach provided a comprehensive analysis of the interplay between cell viability and GFP positive cells under different conditions (**Figure 3 A-C**).

As observed in the three-dimensional representation of Model A, variations in viable cell density and DNA/cell ratio did not significantly affect GFP positive cells (**Figure 3A, B**). Additionally, cell viability was maintained over 80% across all levels of these factors. The non-significant *p* value for cell density in Model B further remonstrated the low relevance of cell density in cell viability (**Table**

2). A significant decrease in cell viability was observed with high PEI concentrations (*p*-value 0.0002, **Table 2**). Of note, when GFP positive cells were predicted to be around 30%, cell viability dropped below 75% (**Figure 3C**). This reveals the impact of PEI concentration on both transfection efficiency and cell viability, indicating that high transfection efficiency in A549 cells is subjected to a remarkable cytotoxic effect.

Therefore, in order to achieve the maximum transfection efficiency while ensuring acceptable cell viability, a combination of Model A and Model B was utilized, with a restriction ensuring cell viability remained above 80%. Under this condition, Model A predicted a maximum transfection efficiency of 18.20%. Without the viability restriction, transfection efficiency increased to 39.19% while viability decreased to 63.05%, which is undesirable given the potential cytotoxic effects of rAAV production-

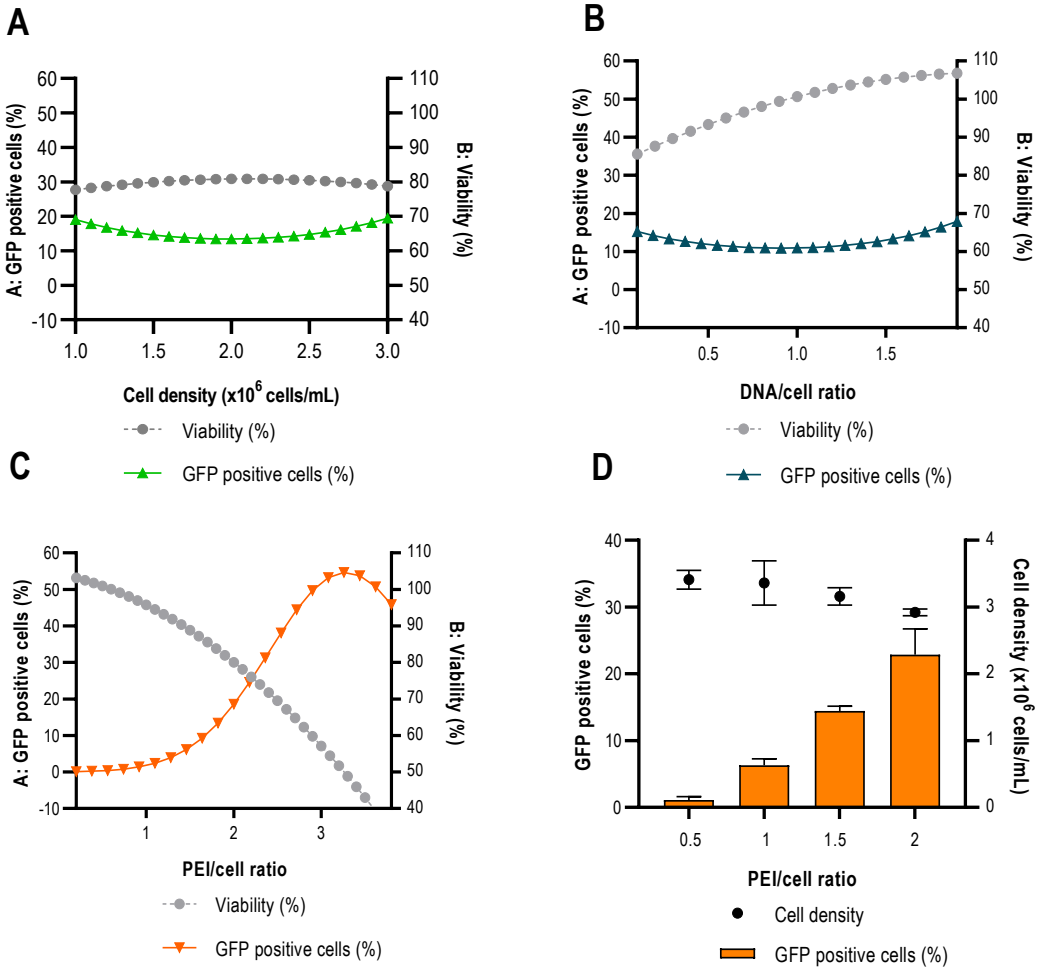


Figure 3. Analysis of variable influences on GFP positive cells (Model A) and cell viability (Model B). A, B, C. The graph depicts predicted values based on Model A (based on GFP positive cells) and Model B (based on cell viability) fixing two of the studied variables and examining the influence on the other. **A.** Analysis of cell density while fixing the optimal values of PEI/cell ratio and DNA/cell ratio. **B.** Analysis of DNA/cell ratio while fixing Cell density and PEI/cell ratio. **C.** Analysis of PEI/cell ratio while fixing DNA/cell ratio and Cell density. **D.** Validation of Box-Behnken design of experiments results analyzing the influence of the PEI/cell ratio variable on GFP positive cells and cell density. DNA/cell ratio was maintained at 1.9 and cell density was fixed at 2.8 x 10^6 cells at the time of transfection. Error bars indicate SD of the triplicates.

related genes. Thus, the optimal conditions for A549 cell line transfection were established as follows: a cell concentration of 2.80 x 10^6 cells/mL, DNA/cell ratio at 1.90 pg/cell and PEI/cell ratio at 2.03 pg/cell. Finally, the predicted optimal values from the combined Model A and Model B were validated. The influence of the PEI/cell ratio was further examined with DNA/cell ratio and cell density fixed at optimal values (1.9 pg/cell and 2.8 x 10^6 cells/mL, respectively) (**Figure 3D**). The highest transfection efficiency of 22% was achieved at a PEI/cell ratio of 2.02 pg/cell, while maintaining a viable cell density at harvest of 85%.

Rational design of E1 construct for constitutive and inducible rAAV production in A549 cells

As a viral vector, efficient rAAV production requires the presence of adenoviral genes, including *E1A, E1B, E2A, E4* and *VA RNAs*[130]. Among them, the *E1A* gene plays a critical role in initiating virus replication by up-regulating transcription from the *rep* gene promoters, p5 and p19[47], and activating the early AdV promoters. Additionally, E1A's role is pivotal in manipulating the cellular environment to promote an optimal cell proliferative state for efficient virus production[101,109,407].

Initially, the feasibility of rAAV production using A549 cells, which are devoid of *E1* gene was addressed, specifically rAAV1 production was analyzed. To this end, E1 functions were supplemented with pE1_cnt plasmid encoding *E1* gene to complement triple transfection (TT) plasmids (namely pGFP_ITR, pHelper and pRep2Cap1) was designed and constructed (**Figure 4A**). E1 control construct was designed based on AdV5 wt genes (AC_000008.1), including nt 194 to 3509 maintaining its endogenous promoters as described in **Materials and Methods** section.

For the construction of the E1A inducible construct, Tet-On® system was selected based on two plasmids. First, this system requires a mutated reverse transactivador, rtTA, which binds to the tetracycline-responsive element (TRE) activating its transcription in the presence of doxycycline (**Figure 4A, B**). Notably, several rtTA variants have been developed to enhance the system performance, reducing basal expression levels while improving inducibility[387,391,392]. Among them, rtTA2^S-M2 was selected, which has demonstrated tight control of gene expression[388]. In our work, prtTA was designed and constructed, containing rtTA^{2}S-M2 alongside the WPRE sequence, known to enhance mRNA stability[393].

The inducible E1A construct, dependent on Tet-On System was devoid of endogenous promoter region (defined as 255 to 501 nt from AdV5 wt genome[256,363,408]), and placed under the control of a TRE (pE1_ind) (**Figure 4A, B**)[391]. The TRE promoter is composed by seven tet-operator sequences (TetO) positioned upstream of the minimal CMV promoter, a shortened version of the

CMV promoter (CMVmin). To note, TetO are the binding sites for the tetracycline-controlled transactivador (rtTA). In pE1_ind plasmid, *E1B* gene is driven by its endogenous promoter, which is positively regulated by E1A protein[101].

Therefore, with the designed Tet-On system, the prtTA transactivador binds to the TRE promoter of the pE1_ind plasmid upon the addition of tetracycline or its derivative, doxycycline[409]. This interaction activates the transcription from the CMVmin promoter downstream gene, driving E1A expression (**Figure 4B**).

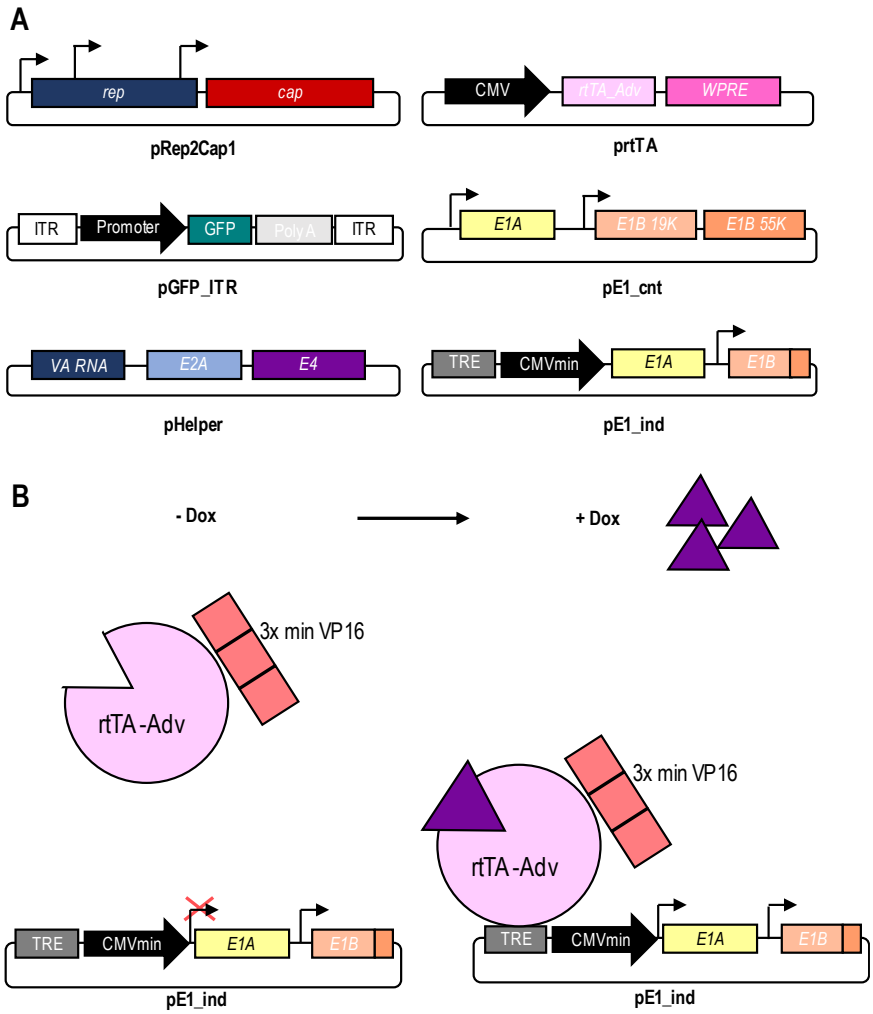


Figure 4. Schematic representation of plasmids designed and constructed. A. The triple transfection (TT) standard plasmids, typically used for HEK293 rAAV production are depicted: pRep2Cap1, pGFP_ITR, and pHelper plasmids. For rAAV production in A549 cells, pE1_cnt is needed, encoding the *E1A* and *E1B* genes. For E1A inducible system, two main components are required: prtTA plasmid containing the rtTA^{2}S-M2 transactivator gene under the control of the CMV promoter with a downstream WPRE sequence and pE1_ind response plasmid, where the *E1A* gene is under control of the tetracycline responsive element (TRE), followed by the CMVmin promoter. **B.** Schematic of the Tet-On Advanced inducible system for E1A expression. Upon the presence of doxycycline (Dox), a tetracycline derivative, the rtTA^{2}S-M2 binds to the TRE, activating E1A expression. Without Dox, E1A remains silent.

Optimization of rAAV1 production in A549 cells

The A549 cell line, lacking the AdV *E1* gene is unable to support rAAV production using only the standard TT plasmids. Consequently, the capacity of the designed and constructed E1 construct (pE1_cnt) to complement the TT for rAAV production was explored. Quadruple transfection was optimized for rAAV1 production using a DoE approach. A four-factor, three-level Box-Behnken experimental design was used to optimize plasmid ratios, namely: pHelper, pRep2Cap1, pGFP_ITR and the pE1_cnt. The experimental design matrix in coded values, response and statistical analysis is presented in **Table 3**. Data were fitted to a quadratic second-order polynomial equation (**Equation 3**) using a linear regression analysis. The generated model coded equation is:

$$Y=512.17+197.51X_1+1.54X_2+249.97X_3+130.45X_4-59.62X_1X_2+88.21X_1X_3+69.07X_1X_4+36.63X_2X_3$$

$$+80.01X_2X_4+106.85X_3X_4-98.01X_1^2-65.04X_2^2-37.56X_3^2-134.46X_4^2 \quad \textbf{(3)}$$

where Y is rAAV1 production (in vg/cell); X_1 is the coded value for pHelper, X_2 is the coded value for pGFP_ITR, X_3 is the coded value for pRep2Cap1 and X_4 is the coded value for pE1_cnt.

Regression analysis showed that the model fitted the data with a coefficient R^2 of 0.9569. This indicates that the model is consistent with 95% of the variability in the data. Further, the statistical significance of the model was confirmed by ANOVA analysis. The Fisher's F-test associated *p* value of <0.0001 indicates that the model is significant. The lack of fit parameter is non-significant (*p* value 0.2286). Values of the tested plasmids pHelper, pRep2Cap1, pE1_cnt are also significant (see **Table 3**).

The second-order polynomial model was used to calculate the optimal plasmid concentrations for rAAV1 production, and to plot response surface graphs (**Figure 5**). The graphs, plotted in three dimensions, illustrate the effects of varying the different plasmid concentration on rAAV1 production (**Figure 5A-F**). Maximal rAAV1 production was achieved near the higher concentration ranges for pHelper, pRep2Cap1 and pE1_cnt. In contrast, the variations in pGFP_ITR concentration did not significantly affect rAAV1 production within the explored design space (*p* value 0.9438).

The optimal plasmid concentrations for maximizing rAAV1 production in A549 cells were determined: 1.24 µg/mL of pHelper, 0.30 µg/mL of pGFP_ITR, 1.36 µg/mL of pRep2Cap1 and 0.74 µg/mL of pE1_cnt. Under these conditions, the predicted rAAV1 production was 1001.27 vg/cell.

Table 3. Matrix design, results, and ANOVA analysis of Box-Behnken experimental design for optimizing rAAV production in A549 cell line by quadruple transfection.

Independent Variables	**Coding Levels**		
	-1	0	1
pHelper (µg/mL)	0.141	0.761	1.381
pGFP (µg/mL)	0.063	0.203	0.343
pRep2Cap1 (µg/mL)	0.201	0.821	1.441
pE1_cnt (µg/mL)	0.023	0.413	0.803

	Variables				**Response**
Experimental Run	**pHelper (µg/mL)**	**pGFP (µg/mL)**	**pRep2Cap1 (µg/mL)**	**pE1_cnt (µg/mL)**	**rAAV1 titer (vg/cell)**
1	0	0	1	-1	286.36
2	0	0	0	0	504.71
3	1	0	0	1	665.56
4	0	0	1	1	724.79
5	0	1	0	-1	152.72
6	-1	1	0	0	98,03
7	0	0	-1	-1	99.94
8	1	-1	0	0	650.19
9	0	1	0	1	618.57
10	0	0	0	0	583.40
11	0	0	0	0	443.81
12	1	0	-1	0	237.24
13	0	0	0	0	545.97
14	0	1	1	0	750.95
15	0	-1	1	0	614.64
16	1	1	0	0	398.00
17	0	-1	0	-1	233.60
18	0	1	-1	0	133.70
19	0	0	-1	1	110.99
20	0	0	0	0	482.95
21	0	-1	-1	0	143.89
22	-1	0	1	0	406.36
23	1	0	0	-1	275.40
24	-1	0	-1	0	27.04
25	-1	0	0	1	148.31
26	-1	0	0	-1	34.32
27	1	0	1	0	969.38
28	-1	-1	0	0	111.74
29	0	-1	0	1	379.41

Model	**Multiple R^2**	***p* value [(a)]**	**Lack of fit test. *p* value [(b)]**
rAAV1 titer (vg/cell)	0.9569	< 0.0001	0.2286
Parameters	**Coeficient**	***f Value***	***p* value [(b)]**
Constant	512.17		
[pHelper]	197.51	84.52	< 0.0001
[pGFP]	1.54	0.0052	0.9438
[pRep2Cap1]	249.97	135.38	< 0.0001
[pE1_cnt]	130.45	36.87	< 0.0001
[pHelper] x [pGFP]	-59.62	2.57	0.1314
[pHelper] x [pRep2Cap1]	88.21	5.62	0.0327
[pHelper] x [pE1_cnt]	69.07	3.45	0.0846
[pGFP] x [pRep2Cap1]	36.63	0.9688	0.3417
[pGFP] x [pE1_cnt]	80.01	4.62	0.0495
[pE1_cnt] x [pRep2Cap1]	106.85	8.24	0.0123
$[pHelper]^2$	-98.01	11.25	0.0047
$[pGFP]^2$	-65.04	4.95	0.043
$[pRep2Cap1]^2$	-37.56	1.65	0.2195
$[pE1_cnt]^2$	-134.46	21.17	0.0004

[(a)] *p values* under 0.05 are considered statistically significant with 95 % confidence, and under 0.1, statistically significant with 90 %.

[(b)] *p values* associated to lack of fit test above 0.05 mean that the hypothesis arguing that the model is suitable cannot be rejected.

Hence, the successful rAAV1 production in the A549 cell line demonstrates the effective design of the pE1_cnt plasmid and its ability to fully support the triple transfection method.

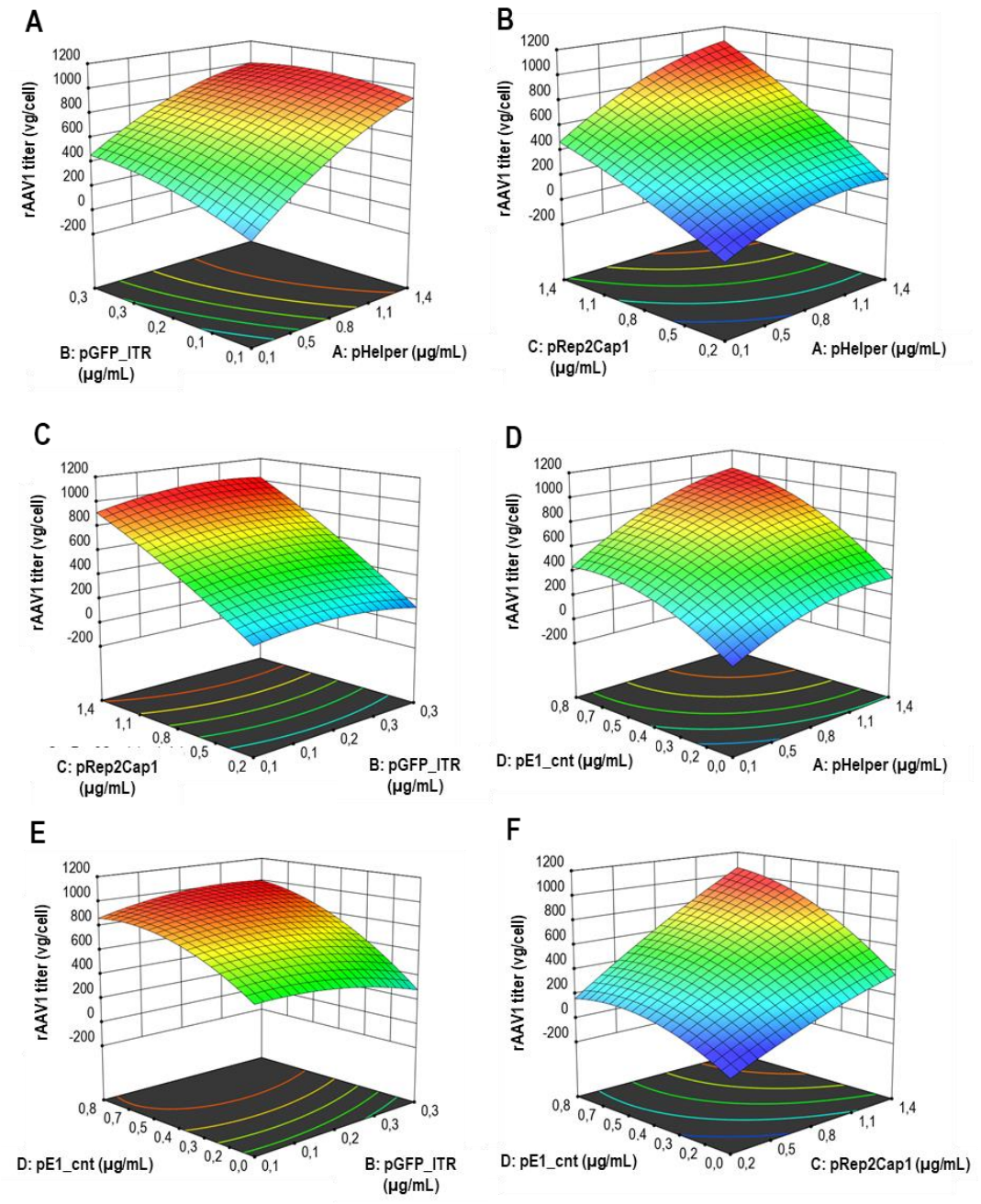


Figure 5. **Response surface graphs based on Box-Behnken design of experiments results for quadruple transfection optimization for rAAV1 production in A549 cells**. The Y-axis accounts for the analyzed response: rAAV1 production (vg/cell). The graphs were constructed by depicting two variables at a time and maintaining the third independent variable at the maximal predicted level: 1.24 µg/mL of pHelper, 0.30 µg/mL of pGFP_ITR, 1.36 µg/mL of pRep2Cap1 and 0.74 µg/mL of pE1_cnt.

Fine-tuning E1A inducible system in A549 cells for efficient rAAV1 production

Having established the optimal conditions for rAAV1 production using an E1 control plasmid in the A549 cell line, the focus shifted towards evaluating the performance of an E1A inducible expression system (represented in **Figure 4**).

Key variables influencing the Tet-On inducible system, such as pE1_ind plasmid and doxycycline dosage on rAAV1 production were investigated. For this, the concentration of pE1_ind plasmid in the transfection mixture was varied within a range from 4.38 x 10^9 to 1.40 x 10^{11} copies/mL, increasing two-fold increase in each condition, while TT and prtTA plasmids were kept constant at 3.5 x 10^{10} copies/mL (**Figure 6A-F**). Doxycycline concentrations ranged from 0 to 2 µg/mL.

The study revealed crucial insights into the balance of E1A gene dosage and inducer concentration for optimal rAAV production. Results indicated that the relationship between pE1_ind dose and rAAV1 production is not linear. Instead, the optimal dose-response effect between doxycycline and rAAV1 titer were observed at the minimal tested pE1_ind dose (**Figure 6A, B**). Increasing doses of E1 inducible construct led to a saturation point, beyond which no further dose-response effect was observed, followed by a decline in rAAV1 titer at highest initial pE1_ind concentrations (**Figure 6D, E, F**). Cell viability remained consistently high across all conditions (**Figure 6A-F**).

Notably, basal expression of the pE1_ind construct without doxycycline was detected, along with rAAV1 production. This leakage showed a proportional relationship with different pE1_ind doses up to 7 x 10^{10} copies/mL (**Figure 6A-F**). This indicates that the E1 construct with inducible promoter can express the E1 proteins at sufficient level to produce rAAV even without doxycycline addition.

Further analysis involved studying E1A expression to determine if the saturation and subsequent decline in rAAV1 production at higher initial pE1_ind doses were due to limitations in the Tet-On system's ability to induce E1A expression effectively. RT-qPCR was conducted, and E1A transcript levels were analyzed across the various pE1_ind plasmid and doxycycline concentrations. As shown in **Figure 6G**, E1A expression was induced at different initial plasmid concentrations, demonstrating that E1A expression itself is not the bottleneck in rAAV production.

Instead, these findings underscore the importance of fine-tuning multiple parameters simultaneously to achieve optimal rAAV1 inducible production, highlighting the complex interplay between gene dosage, inducer concentration, and system performance.

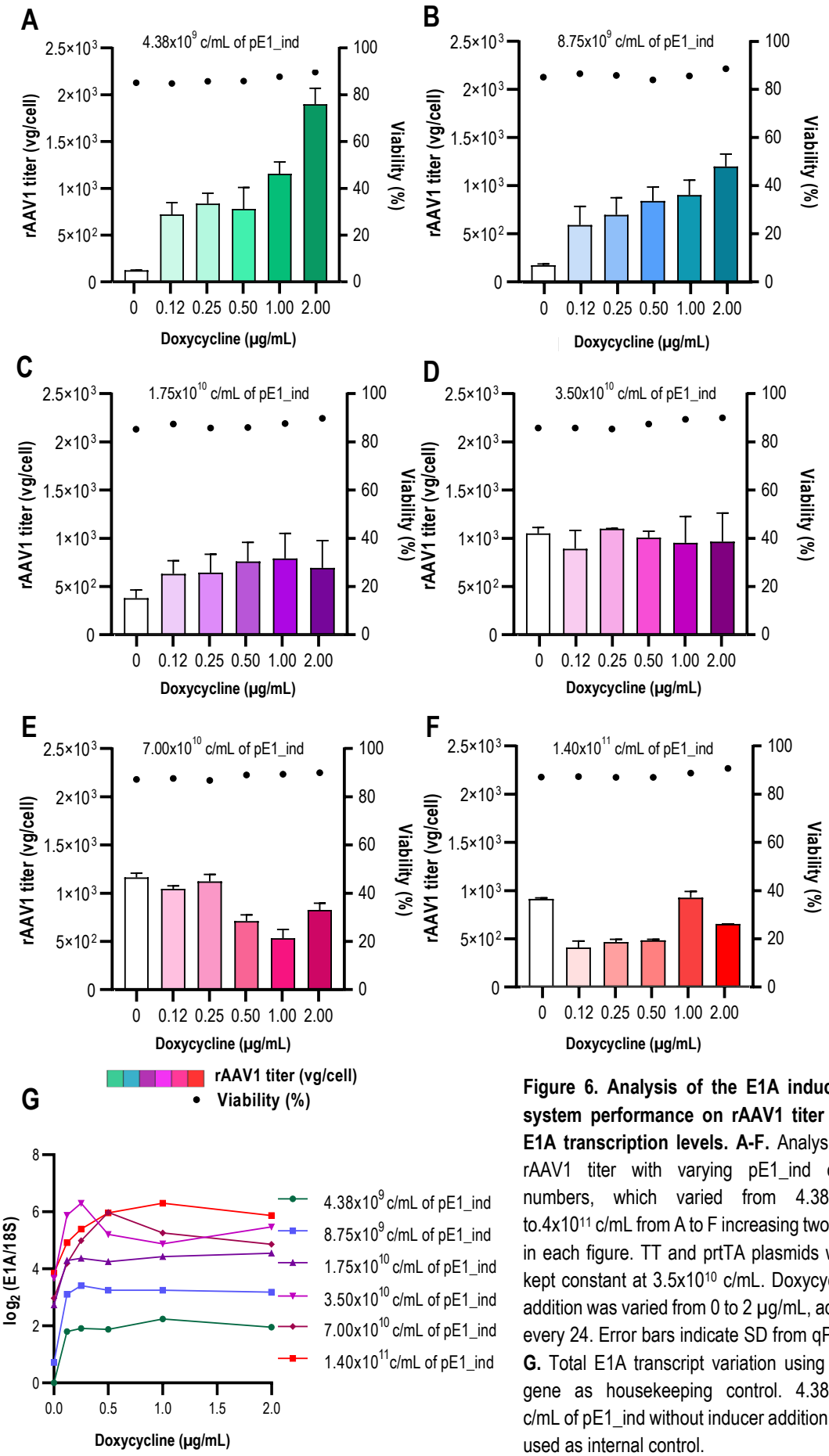


Figure 6. Analysis of the E1A inducible system performance on rAAV1 titer and E1A transcription levels. A-F. Analysis of rAAV1 titer with varying pE1_ind copy numbers, which varied from $4.38x10^9$ to.$4x10^{11}$ c/mL from A to F increasing two-fold in each figure. TT and prtTA plasmids were kept constant at $3.5x10^{10}$ c/mL. Doxycycline addition was varied from 0 to 2 µg/mL, added every 24. Error bars indicate SD from qPCR. **G.** Total E1A transcript variation using 18S gene as housekeeping control. $4.38x10^9$ c/mL of pE1_ind without inducer addition was used as internal control.

rAAV1 production in A549 cells with E1A inducible system is comparable to control system

The quadruple transfection system using pE1_cnt plasmid was compared to the Tet-On-based inducible under optimal conditions to determine the suitability of the inducible system to be integrated into the A549 cell line.

For the control system based on pE1_cnt plasmid quadruple transfection, optimal conditions were previously determined by the Box Behnken design results, reflected in **Equation 3**. The optimal concentrations for the Tet-On-based E1A inducible system were based on the combinations of doxycycline and plasmids leading to maximal rAAV1 production (4.38 x 10^9 c/mL of pE1_ind and 3.5 x 10^{10} c/mL of prtTA, pHelper, pGFP_ITR and pRep2Cap1) shown in **Figure 6A**. In this comparative analysis, rAAV titers were expressed in log-transformed units for relative comparison. Additionally, variations in doxycycline concentrations were analyzed to evaluate their effects on E1A transcript levels and the rAAV1 titer on/off ratio.

Results show that rAAV1 titers obtained with the E1 control system are similar to those achieved by the E1A inducible system with 2 µg/mL of doxycycline addition (**Figure 7A**). Additionally, the inducible system exhibited a log_2 value of 2, indicating a robust dose-response relationship. Further, the addition of doxycycline significantly increased the expression of E1A total transcript as well as its 13S and 12S isoforms (**Figure 7B**)[380]. These findings reveal that similar E1A 13S and 12S transcript expression levels in the inducible system, as compared to the control, correspond to

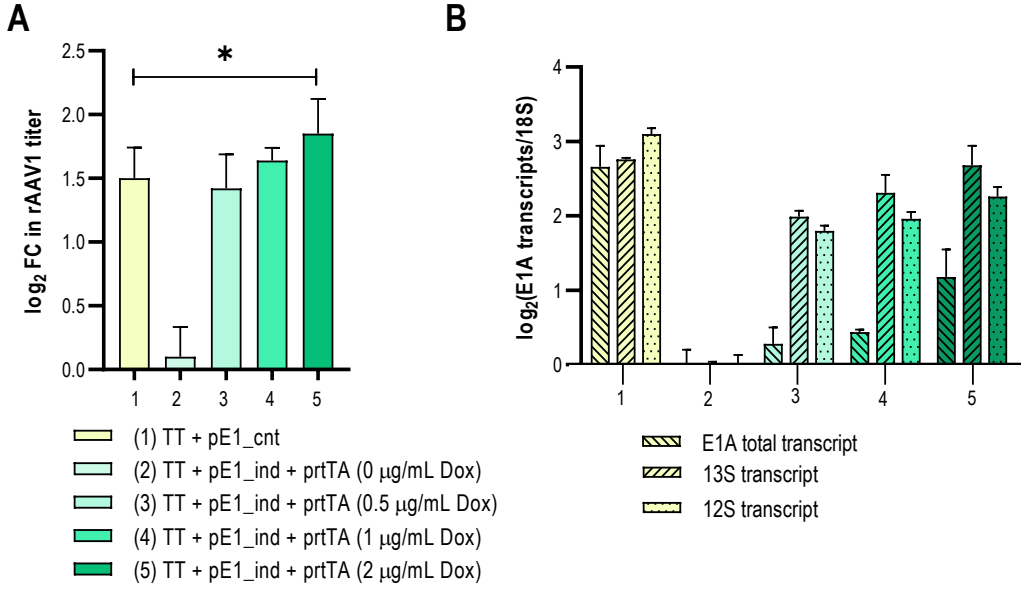


Figure 7. Comparison of rAAV1 production between E1 control system and E1A inducible system in A549 cells. **A.** Log_2FC in rAAV1 titer between control and inducible rAAV1 production systems in A549 cells. **B** Relative transcript levels compared to E1 inducible system without doxycycline addition used as control. Values were normalized to 18S ribosomal gene as housekeeping control. E1A total transcript and isoforms 13S and 12S were analyzed. Data was analyzed by unpaired t-test analysis for significant difference *p<0.05; **p<0.01; ***p<0.001; error bars indicate the SD of biological triplicates.

comparable rAAV titers. This comparative analysis shows that although both the inducible and control systems can yield comparable rAAV1 titers, the inducible system provides significant advantages in terms of modulating E1A expression levels. The capability to precisely control E1A expression via doxycycline addition, which effectively modulates rAAV titer in a dose-responsive manner, is a critical advantage for the development of a stable cell line.

Comparison of rAAV1 quality produced with A549-based and HEK293-based systems

Given that the quality attributes of rAAV preparations can vary depending on the manufacturing platform, biological activity and packaging capacity characterization is critical to demonstrate the production efficiency[369]. This study compared the A549-based platform for rAAV1 production with the current gold-standard, which is transient triple transfection in HEK293 cell line in terms of rAAV1 production and quality. For this comparison, rAAV1 was produced in HEK293 cells under the standard transient transfection conditions using equimolar plasmid ratios of TT plasmids and pMock. In the A549-based system, two production conditions were analyzed: the control E1 expression system and the inducible E1A expression system at the optimal rAAV1 production conditions previously determined.

Table 6. Comparative study of rAAV1 production in A549 and HEK293 cell line. Capsids, vector genome titers and potency of rAAV1-GFP vectors.

Cell platform	Plasmids transfected	Production (vg/mL)	Specific production (vg/cell) [a]	Full capsids (%) [b]	Transducing units (TU/mL) [c]
HEK293	TT + pMock	$8.85x10^9 \pm 1.22x10^9$	3788.09 ± 78.34	17.1 ± 1.00	$1.87x10^8 \pm 7.34x10^6$
A549	TT + pE1_cnt	$2.07x10^9 \pm 2.46x10^8$	923.11 ± 109.60	18.0 ± 1.60	$2.05x10^8 \pm 2.80x10^7$
A549	TT + pE1_ind + prtTA + 2 µg/mL doxycycline	$1.96x10^9 \pm 1.11x10^8$	932.74 ± 52.94	19.2 ± 6.80	$1.90x10^8 \pm 1.92x10^7$

[a] Specific production was calculated using viable cell density at harvest (72hpt).
[b] Full capsids were calculated based on ELISA assay and qPCR determination of full vectors.
[c] Transducing units were calculated using **Equation 2** using three dilutions of rAAV1 purified preparations in triplicate.

While HEK293 cell-based rAAV1 production system displayed enhanced specific production compared to the A549 system (**Table 6**), both systems showed equivalent effectiveness in rAAV1 packaging capacity, achieving similar levels of full capsid across all conditions.

To compare the transduction efficiency of A549 and HEK293-derived rAAV preparations, an in-process compatible purification protocol, adapted from Negreira *et al*.[372] was established, avoiding

the PEG-precipitation step. This adapted protocol enabled low-volume purification while maintaining high-throughput capacity. It also facilitated the purification of triplicate conditions in parallel, enabling a direct comparison of the transduction efficiency of rAAV1 preparations from both systems. Negative TEM analysis confirmed similar levels of purification across all conditions (**Figure 8A**).

The transduction efficiency of the rAAV1-GFP preparations was evaluated using a standard biological assay measuring the expression levels of the GFP transgene contained in rAAV1 vectors. Remarkably, no differences in rAAV1 potency were observed between the two platforms (HEK293 vs. A549) or between the control and inducible systems within the A549 platform (**Table 6** and **Figure 8B, C**). As a result, despite the lower production yields obtained from the A549 cell platform, equal transduction capacity of the generated viral vectors was demonstrated, suggesting that both A549 and HEK293 platforms can produce rAAVs with equivalent potency.

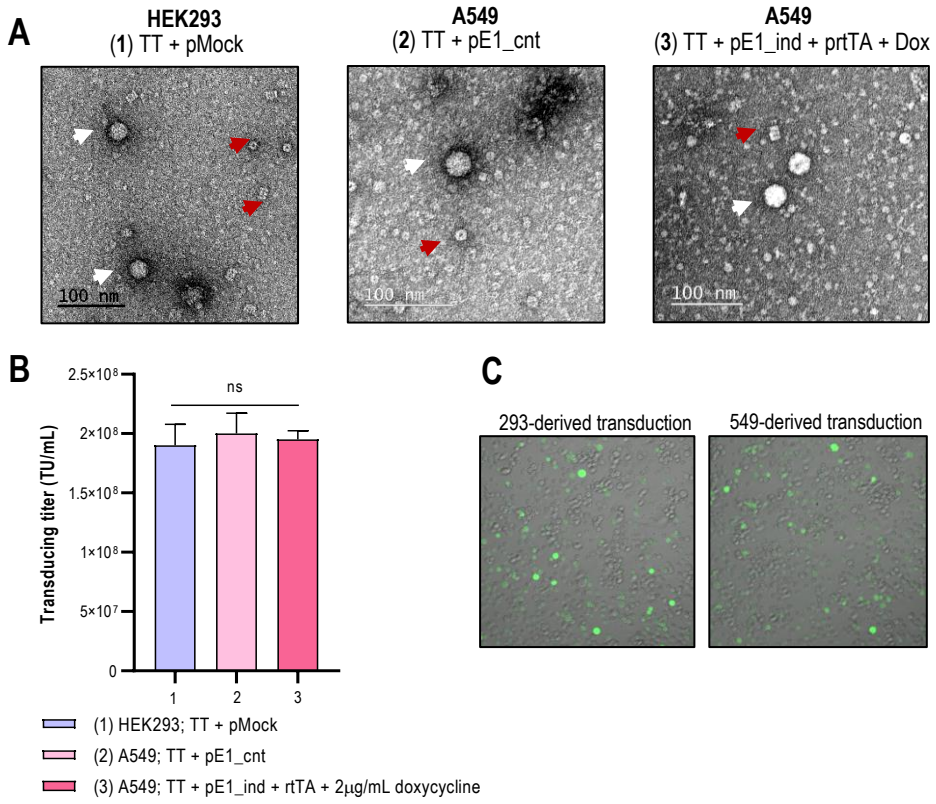


Figure 8. Comparative analysis of rAAV1 transduction efficiency between A549 and HEK293 cell systems. **A.** Negative stain TEM images of purified fractions from different rAAV production systems: HEK293 (1) and A549 constitutive and inducible system, (2) and (3), respectively. White arrows show rAAV particles. Red arrows show the major protein contaminant. **B.** Analysis of transduction efficiency of rAAV1 preparations. HEK293 cells were transduced with rAAV1-GFP purified vector preparations derived from the different cell platforms. GFP positive cells (%) were analyzed by FACS at 48 hours post transduction. Bars represent the mean of TU/mL (**Equation 2**) (±SD) from 3 independent wells and three different infected MOIs. Results were analyzed using One-Way ANOVA. ns, $p>0.05$; *$p<0.05$. **C.** Confocal images of HEK293 cells infected at MOI 177 of rAAV1 preparations from HEK293 and A549 cells.

DISCUSSION

The A549 human lung carcinoma cell line, established in 1972[400], is normally propagated without the need for transformation by AdV *E1A* and *E1B* genes owing to its defect in the p53 and retinoblastoma pathway. This characteristic has favored A549 cells for oncolytic adenovirus production used in cancer therapy[410]. Normally, this cell line is cultured in adherence with serum-containing media. However, scaling up anchorage-dependent cells presents a major challenge. To address this issue, Theriva™ Biologics has developed a serum-free suspension-adapted A549 cell line to streamline large-scale production.

In pursuit of establishing a stable producer cell line for rAAV production, the use of a cell lacking the AdV *E1* gene is identified as a potential strategy to facilitate the integration of all elements for rAAV production. Previous studies have used the A549 cells for the generation of a stable packaging cell line for rAAV production through AAV *rep-cap* gene integration[248]. Although this demonstrated the potential of A549 cell lines for rAAV production also underscored several important limitations[248]. First, even if the *rep* start codon was mutated to express less Rep protein and reduce its toxicity, the process of clone obtention was very challenging, as only 1 out of 1296 clones was successfully obtained[248]. Further, rAAV production required the infection with a temperature sensitive AdV carrying the *E1A* gene, which is undesirable for GMP manufacturing compliance[248]. Finally, although A549 were adapted to suspension culture, the cell culture media required human serum albumin, increasing process costs[248].

The current approach to generate a stable rAAV producer cell line involves utilizing the serum-free suspension A549 cell line developed by Theriva to simplify process scale-up. Additionally, it is proposed to finely tune the E1A gene expression, leveraging its role as a master regulator of rAAV production. The potential of this strategy is evaluated here through the TGE system, maintaining a consistent focus on each step towards the generation of a stable cell line.

A primary limitation encountered in this study was the low transient transfection efficiency in the suspension-adapted A549 cell line, which contrasts with the higher efficiencies seen in traditionally A549-adherent cultures using lipid-mediated systems[411]. Consequently, transfection efficiency was optimized using a DoE approach, tailored to adjust the concentrations of transfection reagent and DNA, achieving optimal transfection efficiency while considering the specific metabolic state of each tested cell density[405]. This optimization process revealed transfection reagent-associated toxicity as the principal factor limiting efficiency, aligning with previous literature[405]. Importantly,

refining the transfection conditions in A549 cells improved efficiency while maintaining sufficient cell viability. Moreover, establishing the optimal for efficient transfection is crucial for stable cell line generation, as enhanced transfection efficiency increases the potential of random integration events.

Progressing towards rAAV production, the pE1_cnt plasmid encoding the E1 proteins was co-transfected with TT plasmids, and rAAV production in A549 cells was subsequently evaluated and quadruple transfection was maximized. Further, a tight and inducible control of the E1A gene expression was obtained using the developed Tet-On inducible system. Characterization of this system revealed a strong influence of the pE1_ind plasmid dosage on rAAV titers. Interestingly, the highest rAAV yield is achieved with the lowest initial E1 concentration tested, exhibiting a proportional dose-response to doxycycline. This suggests that there is an optimal level of E1A expression for maximal rAAV production. This insight is crucial for the successful generation of stable cell lines stably containing inducible *E1A* and *E1B* genes, highlighting the need for precise control over gene expression levels[112].

Other studies have shown that complementing A549 cells with wild type E1A expression reduced the proliferative capacity of A549 cells leading to apoptosis under serum-free conditions[385,412]. In contrast, in the current study, higher E1A expression is not detrimental to A549 cell viability. These findings might suggest that the *E1B* gene anti-apoptotic effect provided in the E1 inducible construct may be beneficial for the rAAV production system.

Notably, the performance of rAAV produced with A549 cell line to that derived from HEK293 cells, the gold-standard for rAAV production, in terms of productivity and potency was compared[293]. As expected, significant differences in the rAAV production yields were observed, which can be attributed to the transfection efficiency levels reached with the two platforms: while HEK293 achieves transfection efficiencies as high as 80%, the optimized transfection efficiency in A549 is around 20%. However, in the context of generating a stable rAAV producer cell line, achieving high rAAV titers in TGE is not the primary objective, but the demonstration that functional rAAVs are obtained. Indeed, the potency of produced rAAV vectors is a critical aspect concerning the choice of cell line for rAAV production. In this study, similar potencies of the produced rAAV vectors, both in A549 and HEK293 cell lines are observed, indicating comparable biological activity. This evidence the potential of using A549 cell line carrying inducible E1A expression as a new approach for the generation of a stable producer cell line for rAAV production.

Relevantly, the tight control of E1A expression shown here has broad potential applications, expanding beyond the A549 cell line. Potentially, stable integration of inducible *E1A* gene could be

implemented in fast-growing cell lines, such as Chinese Hamster Ovary (CHO) cells or Baby Hamster Kindey (BHK) cells, already being considered for rAAV production[413,414]. Moreover, the E1-inducible cells may offer potential for other viral particle production such as first-generation AdV vectors[415,416].

Building upon the effective dose-response rAAV production achieved through E1A induction using the Tet-On system, a significant proof of concept is provided for the generation of a stable producer cell line, particularly using a non-E1 transformed cell line such as A549. This provides the groundwork for the development of a stable A549 cell line with the inducible E1A system stably integrated. Further steps would include the sequential integration of the remaining essential adenoviral helper genes (specifically *E2A*, *E4orf6*, and VA RNA), along with the AAV *rep* and *cap* genes, as well as the rAAV vector transgene sequence. The integration of these components is expected to result in a stable producer cell line capable of producing rAAV in a timely regulated manner triggered by doxycycline addition, thereby eliminating the need for either transient transfection or helper virus infection in rAAV production. The establishment of a cell line harboring all the necessary components for rAAV production would overcome the limitations associated with current transient transfection methods, which are labor-intensive, costly, and often result in variable product quality. Overall, moving one step closer to making gene therapy treatments more accessible.

CONCLUSION

This study successfully optimizes transient transfection in the A549 cell line for rAAV production. Utilizing a Design of Experiments (DoE) approach, including response surface methodology and a Box-Behnken design, optimal conditions for transient transfection are identified, effectively balancing efficiency with cell viability. The rational design of the E1 construct for both constitutive and inducible rAAV production further enhances the potential of the A549 cell line in rAAV production.

The study showcases the potentiality of the A549 cell line, with the developed E1 constructs to drive rAAV production. Despite lower production yields in the A549 platform, the resulting rAAV preparations demonstrate equivalent transduction capacity to those produced in HEK293 cells. This suggests that both platforms are capable of producing rAAVs with similar potencies.

Overall, the results presented here evidence the potential of using A549 cell line carrying inducible E1A expression as a new approach for the generation of a stable producer cell line for rAAV production. This advancement is crucial for scalable and cost-effective rAAV production, significantly enhancing their therapeutic applications.

GENERAL DISCUSSION AND FUTURE DIRECTIONS

Gene therapy products based on recombinant adeno-associated viruses (rAAVs) have received increasing interest in clinical trials and approved applications. However, these products are associated with a high price —often reaching millions of dollars per dose— which significantly limits patient access to these potentially curative treatments. The high costs are partially explained by the fact that current gene therapies address rare diseases with small patient populations, therefore, substantial R&D and production expenses incurred, require companies to set high prices to recoup costs. In addition, the complexity of rAAV manufacturing and the large doses needed per patient contribute significantly to the final price. Consequently, the development of more efficient rAAV manufacturing should be a key factor to contribute to decrease cost of goods and treatment prices, thereby making gene therapy treatments more affordable and accessible. Improved manufacturing processes will not only make gene therapies more viable for rare diseases but also facilitate their expansion to treat more prevalent conditions.

In this work, various strategies aimed at enhancing the rAAV production methods have been investigated. This general discussion provides a summary of the concluding remarks derived from each chapter and proposes how the key findings from this research may serve to future studies focused on the optimization of rAAV manufacturing processes.

In **Part I,** potential improvements for the transient gene expression (TGE) method have been evaluated, specifically through transient transfection into HEK293 cell line for rAAV production. This strategy is nowadays a very extended platform for rAAV production and used for the production of rAAV commercial products, such as Zolgensma®[200].

In previous studies, the Extended Gene Expression (EGE) protocol was developed as a promising tool to improve the rAAV manufacturing process[214]. The enhanced performance of the EGE method, compared to TT for rAAV production has been attributed to its different protocol, combining a Media Exchange (ME) and re-transfection (RT) steps[214]. The ME facilitates the removal of metabolic waste and the supply of fresh media, resulting in increased cell culture viability and enhanced rAAV production[210]. Additionally, the RT step (either one or several) extends the production timeframe by introducing more functional DNA[214]. The use of the EGE protocol for rAAV production has been demonstrated to achieve high rAAV yields in the supernatant, a remarkable advantage not requiring cell lysis to recover AAVs, facilitating the purification process, and yielding comparable amounts to cell lysate recovery in standard batch production[214,295,297]. In addition, the EGE method can be readily transferred to large-scale manufacturing processes through perfusion-based bioprocesses using cell retention devices. This enables the continuous harvest and storage of the supernatant fraction[295,297]. Despite these advancements, the molecular mechanisms

underlying the enhanced efficiency of the EGE method in comparison to the TT for rAAV production remain largely unexplored.

To address this gap, in **Chapter one** a comparative molecular analysis of the effects of different rAAV production methods on HEK293 cells has been performed using a multiplexed quantitative proteomic approach. For this analysis, two rAAV production protocols – the standard TT and the EGE protocol – have been compared to the cell growth condition across different time points. From this comparative study, a total of 4679 proteins have been identified in all three examined conditions. The evaluation of differential protein changes upon the different conditions evaluated within their biological context, has provided crucial insights into the cellular pathways affected during rAAV production. Of note, this study reveals a significant disruption in host cell homeostasis during rAAV production, predominantly characterized by the downregulation in DNA repair and transcription processes and an upregulation in RNA translation pathways as well as lipid biosynthetic route, presumably triggered by the activation of the Unfolded Protein Response (UPR). Of particular interest, the study comparing the EGE method to TT production approach has demonstrated notable metabolic adjustments attributed to the ME and RT steps included in the EGE method, underscoring the capability of these metabolic alterations in the host cell environment to enhance rAAV yield, as extensively discussed in **Chapter one**. These findings offer a valuable perspective for the future rational engineering of host cells aimed at enhancing rAAV production systems, targeting different steps of the transient transfection rAAV production process. Some of the putative targets identified through this analysis have been further explored in **Chapters two** and **three**. However, the potential generated by this proteomic database spans a broad range of applications that can be used in future research on this topic.

In **Chapter two**, the proteomic analysis was used to identify host cell protein targets to increase rAAV production at the extracellular fraction by TT into HEK293 cells focusing on two main steps. In one approach, the capacity to release the rAAVs to the cell supernatant was explored and exploited towards the implementation of improved production strategies allowing to harvest only the supernatant fraction with minimal loss of viral vectors kept inside the cells. Interestingly, it has been demonstrated that the EGE method not only boosts the recovery of rAAV in the cell supernatant by 40% more compared to TT (**Chapter one**), but also rAAV vectors derived from the EGE supernatant fraction are 1.38 times more potent than those harvested from TT (**Chapter two**). This also correlates with an increased recovery of rAAV post-ultracentrifugation and a higher EV harvest, indicating a higher ratio of EV/vg in EGE compared to TT production method. These

findings suggest that the EGE operation mode is associated with higher EV co-production alongside rAAVs.

Indeed, previous studies reported that a portion of rAAV vectors are naturally released associated with extracellular vesicles (EVs) into the supernatant of the cell culture media, named as exosome-associated rAAVs (exo-rAAV)[151,152,345]. These exo-rAAVs have received interest due to their superior transduction properties compared to conventional rAAV vectors, resistance to neutralizing antibodies[346–348], and higher efficiency in transducing hard-reaching tissues crossing the blood-brain barrier[155] or the retina's inner membrane[349]. Thus, the use of the EGE method not only outperforms the TT in rAAV yields, but it can facilitate the production of exo-rAAVs as a valuable pharmaceutical product[152,346–348]. The generation of rAAVs with enhanced potency as the exo-rAAVs, is a critical attribute, enabling the administration of lower rAAV doses for the same clinical effect, potentially reducing the immunogenic reactions associated with viral vector toxicity.

Building on these findings, one of the mechanisms contributing to the viral egress in the EGE method is considered to be in association with EV release, with rAAV particles being secreted from the cytoplasm via exosomes[156]. This hypothesis aligns with prior findings, including the AAV membrane-associated accessory protein (MAAP), identified as a viral egress factor that facilitates rAAV exit from cells by interacting with the outer surface of EVs[89]. In consequence, the thesis has explored how to improve the exo-rAAV egress pathway to increase the proportion of rAAVs exported to the extracellular fraction. The proteomic study in **Chapter one,** has revealed that vesicle transport pathways are upregulated in EGE compared to the TT method of rAAV production. Notably, key metabolic targets within the pathways of MVB loading and exosome release, particularly those overexpressed in the EGE method compared to TT, have been identified. This candidate proteins, including SNAP47, VPS37B, SMPD3 and CD63, have proved effective as modulators of the exosome pathway in mammalian cells, such as HEK293 and CHO-S. Moreover, their overexpression through a host cell metabolic engineering approach has led to an increase in extracellular rAAV9 titer and egress rate at various levels during TT in HEK293 cells. In this work such protein candidates have been overexpressed using plasmids. Future work generating stable cell lines with the identified protein candidates overexpressed could be valuable serving dual purposes: characterizing the obtained rAAV products and maximizing the extracellular rAAV yield produced by transient transfection into the engineered HEK293 cell line with improved metabolic characteristics.

First, the generation of stable cell lines with identified candidates in **Chapter two**, particularly with CD63 overexpression, could help to characterize the association between rAAVs and EVs. This

involves a detailed analysis of the proportion of rAAVs either associated with or encapsulated within EVs, and assessing how this association affects the transduction efficiency. Notably, it has been demonstrated that the rAAV-containing exosomes, devoid of free-rAAVs, facilitate cell uptake and entry while offering protection against antibody neutralization[417]. Therefore, investigating how each targeted protein influences the encapsulation of rAAVs within EVs, yet unexplored in this work, could hold significant implications for rAAV in clinical applications. Second, the generation of stable cell lines incorporating a combination of VPS37B and SMPD3, identified in **Chapter two** for providing the highest rAAV yields in the extracellular fraction, could be of use not only to generate high quantity of free rAAV particles in the cell supernatant but also to explore improvements in their quality characteristics, and future work should be pursued in this direction.

In the scope of the generation of higher quality rAAV vectors, previous studies conducted by Vandenberghe *et al*., showed that rAAVs collected from different cell fractions exhibited similar potency in experiments conducted with adherent HEK293 cells in serum-containing media[151]. Interestingly, observations on Parvovirus infectivity studies suggested that Parvoviruses mature and gain full infectivity as they progress through the secretory pathway[333]. Consequently, the modification of the rAAV secretory pathway through the EV-targeting protein candidates identified in this work, could lead to the production of more mature rAAVs, potentially leading to the creation of vectors of superior quality. Building on this concept, several proteins identified in the proteomic study detailed in **Chapter one** could be leveraged to enhance rAAV quality by targeting pathways that have not been analyzed in this work nor previously reported in the literature discussed hereafter.

The proteomic study showed that the TT rAAV production method was associated with significant cellular homeostasis disruption characterized by the activation of the UPR response pathways. The UPR is a cellular stress pathway activated in response to the accumulation of unfolded or misfolded proteins in the ER lumen with the purpose of restoring the normal function by stopping the protein translation, degrading misfolded proteins, and activating the signalling pathways that increase the production of molecular chaperones involved in protein folding. Excessive ER stress could also lead to apoptosis halting rAAV production[319]. In contrast, the EGE method, compared to the TT is associated with an enhanced cell homeostasis reflected by increased cell viability at harvest. At the proteomic level, the EGE production method could modulate the host cell metabolism to counteract cell homeostasis disruption and mitigating ER stress, which is highly induced by rAAV production. Remarkably, the accumulation of misfolded capsid proteins in the cytoplasm could impact rAAV production. Indeed, a proteomic analysis of rAAV production reported by Lu *et al*.

revealed that capsid proteins undergo proteasome degradation by host cell response, as after adding a proteasome inhibitor, higher level of AAV2 capsid protein was obtained, boosting rAAV2 production[418]. Contrarily, the proteomic study conducted in this work suggests that while the misfolded accumulation of capsid proteins may be a bottleneck for rAAV production, halting protein degradation may not be the most effective approach. Instead, a strategy that mimics the pathways activated in the EGE method, which are aimed at modulating host cell homeostasis to reduce the UPR response, could be more beneficial. Accordingly, it is proposed that the overexpression of cytosolic chaperones, identified in this study upregulated in EGE compared to TT, such as HSPA1 or ubiquitin-conjugating enzymes might support capsid protein folding or prevent the accumulation of misfolded proteins by degrading them. This approach could lead to an improvement in rAAV capsid assembly quality and quantity[40].

Furthermore, in this study, KPNA2, a component of the nuclear pore complex (NPC) involved in the nuclear import of rAAV, was highly upregulated in the EGE method of rAAV production. This suggests that the EGE protocol may favour the nuclear entry of rAAV capsid proteins. We hypothesize that HEK293 cell overexpression of KNPA2 basal levels may increase availability of rAAV capsids in the nucleus which may potentially lead to increased encapsidation efficiency.

The analysis of the proposed targets may guide future host cell engineering efforts to enhance rAAV production via transient transfection, alongside improving the quality of the produced viral particles.

Additionally, one of the challenges limiting large scale implementation of TGE for batch processes is the cell density effect (CDE), which is the reduction in transfection efficiency at cell concentrations typically over 2 x 10^6 cells/mL, beyond the mid-exponential growth phase[281]. To address this issue, considering the results obtained in **Chapter one**, media supplementation feeds have been developed, formulated with fatty acid and cholesterol concentrates with the aim to counteract the decrease in lipid biosynthetic pathways observed with increased cell densities. Adding this lipid supplementation before transfecting the cells at 6 x 10^6 cells/mL, transfection efficiency improved up to 50%, as opposed to the 30% efficiency observed in the absence of lipid supplementation. Overall, the lipid addition protocol developed in **Chapter two**, proved to be a potential strategy towards scalable TT of HEK293 cells.

In **Chapter three** the shift has been focused into adenoviral *E1* gene expression alterations during rAAV production within HEK293 cells. These findings have substantial implications for rAAV production methodologies, whether they rely on transient transfection or stable cell lines. Importantly, it has been evidenced that during the rAAV production process via transient

transfection in HEK293 cells, there is a downregulation on endogenous E1A expression levels compared to cell growth condition. Given that E1A protein is pivotal in initiating the rAAV production process being considered its master regulator, its downregulation could have implications in rAAV production yields. In this work, it has been observed that the E1A downregulation is triggered by the expression of adenoviral VA RNAs or the E4 gene ORF. As both are essential for rAAV production, this is a significant discovery with potential impact on rAAV production. To the best of our knowledge, this interaction has not been previously documented and thus opens new avenues for research on rAAV production dynamics.

Several strategies have been explored to mitigate the challenge of reduced intrinsic E1A expression during rAAV production in HEK293 cells. In one strategy, a ME step, performed 48 hours post-transfection, with or without a subsequent re-transfection (as in the EGE protocol) has been found to preserve the E1A expression levels comparable to those observed under standard cell growth conditions. Further, this significantly boosted rAAV production and viral egress. The ME strategy, as discussed, can be implemented in continuous manufacturing bioprocesses by perfusion operation mode using cell retention devices. The cell specific perfusion rate (CSPR) is indicative of the medium exchange rate per cell per day in a perfusion process and is cell line and medium dependent. Building on the findings of this research, it is suggested that investigating the preservation of E1A expression and its subsequent impact on rAAV production varying the CSPR, could maximize rAAV productivity. Furthermore, a multi-omic study based on such experiments could generate insights into energetic demands of cells during continuous media exchange. Alternatively, the use of fed-batch and its impact on E1A expression, not been explored in this work, could represent a simpler alternative to continuous media exchange, if effective. Though it limits the continuous viral harvest capability compared to perfusion methods, it may preserve E1A expression, potentially expanding the system productivity without requiring cell retention devices thereby reducing process complexity.

Noteworthy, the perfusion-based production system has been effectively implemented for rAAV production both through the TT method into HEK293 cells and using stable rAAV producer cell lines such as ELEVECTA™, developed by CEVEC Pharmaceuticals[210,339,419]. Noteworthy, extensive efforts have been directed towards the development of rAAV producer stable cell lines for stable gene expression (SGE) production method. However, such stable cell lines with all required elements for rAAV production in the context of the results that are provided here, reveal that the E4 ORF and VA RNAs not only negatively impact cell viability, as previously documented in the literature[198,374], but also have an inhibitory effect on E1A expression.

In these rAAV stable producer cell lines, perfusion-based approaches could be instrumental in preserving E1A expression. Perfusion methodologies have been reported to offer multiple advantages in the rAAV production context[419]. Not only they increased cell-specific virus yield compared to conventional batch process but also dramatically increased the ratio between full and empty viral particles by a 7 to 9-fold in perfusion-based compared to the traditional batch process[419]. This suggests that investigating the molecular changes on host cells exerted by ME could give insights into enhancing rAAV productivity, and importantly, packaging efficiency.

The second method explored in **Chapter three** to prevent the E1A expression decrease during rAAV production has been plasmid co-transfection encoding the adenoviral *E1* gene ORF. This approach significantly enhanced rAAV yield, surpassing the ME strategy in both the cell supernatant and cell lysate fractions for rAAV1, rAAV6, and rAAV9. For rAAV8, although the metabolic engineering approach yielded higher specific production rates in both the extracellular and cell lysate fractions, it did not outperform the ME strategy in the cell supernatant. However, using the E1 gene cassette raises safety concerns. During rAAV production, DNA sequences can be encapsidated into AAV particles in a mechanism that might be mediated by the promiscuous binding of the AAV Rep protein to RBE-like motifs in DNA sequences, leading to plasmid DNA packaging up to 1 to 3%[220,420]. The use of an E1A-containing plasmid, particularly when transfected in excess, could increase the risk of rAAV vectors to encapsulate this oncoviral sequence. In future work, the evaluation of the pseudo-packaging efficiency of E1A, both supplied by the plasmid and the by the E1 cassette present in HEK293 cells is proposed.

Furthermore, in this proposed system the addition of an additional plasmid to the TT system adds additional cost to the process. Future efforts based on this work should aim to replicate the effects of E1A preservation or overexpression in HEK293 through indirect strategies leading to E1A-like activities. These approaches could be based on the multiple roles of E1A in the rAAV production context. One avenue involves mimicking the E1A's role on cell cycle progression. E1A facilitates the transition from the G0 or G1 into the S phase of the cell cycle, creating an optimal environment for virus replication. This is achieved by inactivating the pRb checkpoint, thereby freeing the E2F transcription factor to activate genes essential for nucleotide metabolism and DNA replication required in the S phase[324,325]. Thus, E1A activity could be simulated by expressing factors that promote entry into the S phase or inhibit those that restrict S-phase progression, thereby favouring the G1/S phase transition. The overexpression of Cyclin E, which activates E2F and interacts with Cdk2 to drive cell cycle progression from G1 to S phase, might be effective for this purpose. Further, E1A also has a significant role in interacting with transcriptional modulators, including histone

acetyltransferases (HATs), histone deacetylases (HDACs), and other chromatin remodelling factors, influencing cellular and viral gene transcription[326]. In this context, chemical compounds altering chromatin accessibility for transcription could be used such as Trichostatin A, which is a potent HDAC inhibitor that increases histone acetylation, thereby promoting a more open chromatin structure[330,421].

Overall, in **Part I**, various strategies have been successfully implemented to improve the rAAV production method through standard TT in HEK293 cell lines. While the TT method is well-suited for research and early clinical trials as it generates vectors readily, scaling it up for larger-scale production implies significant challenges, including high cost of goods and high process variability, inherent to the transient transfection process[194]. To address these challenges, **Part II** explores the use of inducible systems towards a proof of concept for the generation of stable rAAV producer cell line, aiming to a more robust process and enhanced control over product quality as an alternative to transient transfection systems[247]. To this end, the choice of the cell line is key and different alternatives have been explored in this work as discussed in **Chapters four** and **five**.

In **Chapter four,** the feasibility of using the HEK293 cell line towards the development of a stable rAAV packaging cell line has been investigated. However, the use of HEK293 cells for generating packaging cell lines for rAAV production, containing *rep-cap* genes integrated is particularly challenging as it constitutively express *E1A* and *E1B* genes from AdV[181], which leads to constant E1A-mediated transactivation of the AAV p5 promoter that controls the expression of the cytotoxic Rep proteins[58,104,197]. To address the Rep toxic expression, a stepwise rational design approach has been employed. The four different Rep proteins have been individually expressed under the inducible Tet-On system and employing a Design of Experiments (DoE) methodology, the optimal conditions for a *rep*-inducible system that maximizes rAAV production have been explored. A Rep construct with a proportional dose-response between the inducer and rAAV production has been identified, with minimal toxicity at the non-induced state. However, the rAAV production efficiency remained low. This limitation may be due to the use of non-endogenous promoters, which may be detrimental to rAAV production failing to provide the precise temporal regulation necessary for a rAAV production system.

Nonetheless, the development of a stable cell line carrying the Rep-inducible constructs developed in this work may be of interest for its ability of exogenously control Rep protein expression. Importantly, the coordinated expression of the replicated rAAV genome and capsid proteins has been shown to influence packaging efficiency[422]. Thus, the development of a *rep*-inducible HEK293 stable cell line, capable of timely control *rep* gene expression while maintaining capsid gene

expression at high levels consistently by transient transfection, could be used for improving encapsidation efficiency[275,422].

Given the challenges identified with stable rAAV producer cell lines in **Chapter four**, **Chapter five** focus has been shifted to the A549 suspension serum-free cell line, which lacks E1 adenovirus expression. Notably, previous work addressing the stable rAAV producer cell line generation using HeLa cells, which are also devoid of E1 expression, was unsuccessful[257]. This attempt used the Tet-Off system which is non-desirable for large scale implementation, and E1A inducibility potential was non-stable over time[257]. In contrast, this thesis represents a first effort towards a stable rAAV producer cell line development using A549 cell line by E1A inducible expression using the Tet-On system. Key effectors influencing the E1A inducible expression system have been evaluated in A549 cells, impacting rAAV production towards a stable cell line development based on this platform. Despite lower production yields in the A549 platform compared to the current gold standard HEK293 TT system, attributable to the low transfection efficiency of the A549 cell line and the requirement of more plasmids for rAAV production, the resulting rAAV vectors show equivalent transduction capacity to those produced in HEK293 cells. These results underscore the potentiality of the A549 cell line as a novel rAAV production platform.

Although shifting from HEK293 cells, which inherently express the adenoviral *E1* gene, to A549 cells might seem like a step back, the precise exogenous control of E1A expression presented here offers a strategic advantage for rAAV production, also offering a potential to circumvent the limitation of E1A decrease triggered by adenoviral genes as documented in **Chapter three**. Relevantly, the tight control of E1A expression achieved in this study has potential applications that could expand beyond A549 cell line. Stable integration of inducible E1A gene as the system developed here could be implemented in fast-growing cell lines, such as Chinese Hamster Ovary (CHO) cells or Baby Hamster Kindey (BHK) cells, which are commonly used in industrial settings and an emerging platform currently explored for rAAV production[413,414]. Further, the E1-inducible cell line, may offer potential for other production uses such as first-generation AdV vectors[415].

Building on the progress achieved in this work, the generation of a stable rAAV producer cell line based on the A549 cell line could offer an effective alternative to the ones developed to date[274,275]. To this end, a stepwise approach is proposed aiming to generate a stable rAAV production platform, initially by the integration of the developed E1A inducible system, followed by the stable introduction of the essential adenoviral genes necessary for rAAV production and finally, the stable integration of the replication and packaging module. Noteworthy, in this approach the replication and packaging module may be specifically modified aiming to generate a versatile cellular platform

adaptable to produce any rAAV product of interest. In this regard, despite the challenges encountered in **Chapter four** with the HEK293 cell line, the Rep-cap split strategy designed and implemented emerges as a promising solution for flexible cell line development. The independent expression of the *cap* gene driven by a strong constitutive promoter enables its flanking with a recombinase-recognition target sequences, that would enable the possibility for the capsid sequence to be readily exchanged through Recombinase-Mediated Cassette Exchange (RMCE) method. This serotype-switching capability has not been approached before in rAAV producer cell line generation and adds interest in this platform, as for any new rAAV serotype to be produced, stable cell line can be flexibly modified using the RMCE, which can be performed within weeks.

Noteworthy, the contributions from **Part I** may be of relevance towards enhancing the productivity and quality of rAAV vectors, applicable not only in TT rAAV production using HEK293 cells, but also may be of use in developing a stable rAAV producer cell line based on A549 cells.

Overall, this study establishes the groundwork for the development of scalable and robust rAAV production systems capable of producing high quantity and quality vectors. These developments hold the promise to significantly impact gene therapy field, offering cost-effective processes for producing rAAV-based treatments, ultimately improving patient accessibility.

CONCLUSIONS

From the results obtained in this work, the following conclusions can be highlighted:

Part I: Enhancement of rAAV production through transient transfection

1. The transient transfection (TT) method of rAAV production triggers a significant disruption in HEK293 cell homeostasis, reducing cellular viability, characterized by the downregulation in DNA repair and transcription processes and an upregulation in RNA translation pathways as well as lipid biosynthetic pathways. Overall, the activation of the Unfolded Protein Response is observed.

2. The extended gene expression (EGE) method compared to the standard TT production approach, demonstrates notable metabolic adjustments attributable to Media Exchange (ME) and Re-Transfection (RT) steps. These differences include modulation of metabolic processes rendering a reduced mitochondrial activity, differential cell cycle modulation, upregulation of lipid biosynthetic processes and enhanced vesicle transport pathways.

3. The EGE method enhances rAAV specific production and viral egress compared to the TT approach. The ME step in the EGE is identified as the step leading to enhanced viral egress, improving from 38 to 44% for rAAV9, and 8 to 15% for rAAV1.

4. rAAV9 particles isolated from the extracellular fraction of the EGE method exhibit enhanced association with extracellular vesicles (EVs) and 1.38 times more potency compared to those produced by the TT method.

5. Overexpression of CD63, VPS37B, SNAP47, and SMPD3, increase EV production in HEK293 cells and remarkably in CHO-S cells (e.g $1.5x10^{10}$ EV/mL with CD63 overexpression compared to $3.5x10^{9}$ EV/mL in control condition). Maximal egress of rAAV9 was achieved when SMPD3 and VPS37B are combined at ratios of 0.415 and 0.585, respectively, achieving a 50% egress rate.

6. Media supplementation feeds formulated with cholesterol added prior to transfecting cells at a high cell density of 6×10^{6} cells/mL increase transfection efficiency up to 50% compared to the 30% observed without supplementation.

7. The downregulation of E1A expression during rAAV production by TT is triggered by the adenoviral helper plasmid expression, specifically E4 ORF and VA RNAs, but not E2A ORF.

8. E1A protein preservation (by ME) or overexpression (by metabolic engineering) enhances rAAV production across multiple serotypes. For rAAV1, rAAV6, rAAV9, metabolic engineering with E1 supplementation yields higher than ME in both supernatant and lysate fractions. For rAAV8, metabolic engineering increases titers but does not outperform ME in supernatant.

Part II: Proof of concept of an inducible system towards stable rAAV cell line generation

9. Rep-inducible expression through the Tet-On inducible system is highly dependent on initial concentrations of *rep* and inducer addition. The regulation of this parameters leads to a proportional dose-response range between inducer and rAAV1 production.

10. In the developed Tet-on system, large Rep proteins and not small Rep proteins are essential for rAAV production. The pRep68 and pRep78mutS constructs demonstrate the maximal rAAV production, higher encapsidation rates and lower toxicity under non-inducible conditions, thereby identified as promising candidates towards a stable packaging cell line generation in HEK293 cells.

11. Transient transfection efficiency maintaining cell viability was optimized in A549 cell line. Optimal conditions were determined: cell concentration of 2.80 x 10^6 cells/mL, DNA/cell ratio at 1.90 pg/cell and PEI/cell ratio at 2.03 pg/cell. Using these conditions transfection efficiency reached 22%.

12. The optimal plasmid concentrations for maximizing rAAV1 production in A549 cells were determined: 1.24 µg/mL of pHelper, 0.30 µg/mL of pGFP, 1.36 µg/mL of pRep2Cap1 and 0.74 µg/mL of pE1_cnt. Under these conditions, the predicted rAAV1 production was 1001.27 vg/cell.

13. The A549 cell line for rAAV production using the inducible E1 expression system, presents similar quality and potency characteristics compared to those produced by TT into HEK293 cells (1.87x10^8 ± 7.34 x 10^6 TU/mL in rAAV1 produced by TT into HEK293 cells compared to 1.90 x 10^8± 1.92 x 10^7 TU/mL under the maximal obtained conditions for A549-based E1A inducible system).

REFERENCES

1. Directive 2001/83/EC, Annex I, Part IV, as amended in Directive 2009/120/EC.
2. Arabi, F., Mansouri, V. & Ahmadbeigi, N. Gene therapy clinical trials, where do we go? An overview. *Biomed. Pharmacother.* **153**, 1–20 (2022).
3. Au, H. K. E., Isalan, M. & Mielcarek, M. Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings. *Front. Med.* **8**, 1–14 (2022).
4. Sheridan, C. Gene therapy finds its niche. *Nat. Biotechnol.* **29**, 121–129 (2011).
5. Williams, D. A., Lemischka, I. R., Nathan, D. G. & Mulligan, R. C. Introduction of new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature* **310**, 476–480 (1984).
6. Friedmann, T. & Roblin, R. Gene Therapy for Human Genetic Disease? *Science.* **175**, 949–955 (1972).
7. Sambrook, J. *et al.* T Lymphocyte-Directed Gene Therapy for ADA-SCID: Initial Trial Results After 4 Years. *Science.* **270**, 475–480 (1995).
8. Kaufmann, K. B., Büning, H., Galy, A., Schambach, A. & Grez, M. Gene therapy on the move. *EMBO Mol. Med.* **5**, 1642 (2013).
9. Puumalainen, A. M. *et al.* Beta-galactosidase gene transfer to human malignant glioma in vivo using replication-deficient retroviruses and adenoviruses. *Hum. Gene Ther.* **9**, 1769–1774 (1998).
10. Bordignon, C. *et al.* Gene Therapy in Peripheral Blood Lymphocytes and Bone Marrow for ADA − Immunodeficient Patients . *Science.* **270**, 470–475 (1995).
11. Stolberg, S. G. The biotech death of Jesse Gelsinger. *N Y Times Mag.* **28**, 149–150 (1999).
12. 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).
13. Wilson, J. M. Lessons learned from the gene therapy trial for ornithine transcarbamylase deficiency. *Mol. Genet. Metab.* **96**, 151–157 (2009).
14. Cavazzana-Calvo, M. *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science.* **288**, 669–672 (2000).
15. Hacein-Bey-Abina, S. *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* **118**, 3132–3142 (2008).
16. Ylä-Herttuala, S. Endgame: Glybera finally recommended for approval as the first gene therapy drug in the European union. *Mol. Ther.* **20**, 1831–1832 (2012).
17. Pharmaphorum. Glybera, the most expensive drug in the world, to be withdrawn after commercial flop. (2017). Available at: https://pharmaphorum.com/news/glybera-expensive-

drug-world-withdrawn-commercial-flop/. (Accessed: 7th March 2024)

18. Aiuti, A., Roncarolo, M. G. & Naldini, L. Gene therapy for ADA-SCID, the first marketing approval of an ex vivo gene therapy in Europe: paving the road for the next generation of advanced therapy medicinal products . *EMBO Mol. Med.* **9**, 737–740 (2017).
19. Padhy, S. K., Takkar, B., Narayanan, R., Venkatesh, P. & Jalali, S. Voretigene neparvovec and gene therapy for leber's congenital amaurosis: Review of evidence to date. *Appl. Clin. Genet.* **13**, 179–208 (2020).
20. Paul-Ehrlich-Institut. Gene Therapy Medicinal Products. (2024). Available at: https://www.pei.de/EN/medicinal-products/atmp/gene-therapy-medicinal-products/gene-therapy-node.html?cms_gts=166822_list%253DdateOfIssue_dt%252Basc. (Accessed: 1st March 2024)
21. Heo, Y. A. Etranacogene Dezaparvovec: First Approval. *Drugs* **83**, 347–352 (2023).
22. Pagliarulo, N. Pricey new gene therapies for sickle cell pose access test. (2023). Available at: https://www.biopharmadive.com/news/crispr-sickle-cell-price-millions-gene-therapy-vertex-bluebird/702066/. (Accessed: 11th May 2024)
23. De Luca, M. & Cossu, G. Cost and availability of novel cell and gene therapies. *EMBO Rep.* **24**, 2–5 (2023).
24. Wirth, T., Parker, N. & Ylä-Herttuala, S. History of gene therapy. *Gene* **525**, 162–169 (2013).
25. European Medicines Agency (EMA) 'CAT Quarterly Highlights and ATMP Approvals'. (2023). Available at: https://www.ema.europa.eu/en/documents/report/cat-quarterly-highlights-approved-atmps-july-2023_en.pdf. (Accessed: 1st March 2024)
26. Ginn, S. L., Amaya, A. K., Alexander, I. E., Edelstein, M. & Abedi, M. R. Gene therapy clinical trials worldwide to 2017: An update. *J. Gene Med.* **20**, 1–51 (2018).
27. Dan Wang, G. G. State-of-the-art human gene therapy: part I. Gene delivery technologies. *Discov Med* **18**, 67–77 (2014).
28. Wang, C. *et al.* Emerging non-viral vectors for gene delivery. *J. Nanobiotechnology* **21**, 1–18 (2023).
29. Ramamoorth, M. & Narvekar, A. Non viral vectors in gene therapy - An overview. *J. Clin. Diagnostic Res.* **9**, 1–6 (2015).
30. Thomas, C. E., Ehrhardt, A. & Kay, M. A. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4**, 346–358 (2003).
31. Cabanes Creus, M. Novel AAV engineering technology: identification of improved AAV variants for gene addition and genome engineering in primary human cells. *Doctoral thesis, UCL (University College London)*. (2019).

32. Bouard, D., Alazard-Dany, N. & Cosset, F. L. Viral vectors: From virology to transgene expression. *Br. J. Pharmacol.* **157**, 153–165 (2009).

33. Atchison, R. W., Casto, B. C. & Hammon, W. M. Adenovirus-associated defective virus particles. *Science.* **149**, 754–756 (1965).

34. Berns, K. I., Pinkerton, T. C., Thomas, G. F. & Hoggan, M. D. Detection of adeno-associated virus (AAV)-specific nucleotide sequences in DNA isolated from latently infected Detroit 6 cells. *Virology* **68**, 556–560 (1975).

35. 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**, 2077–2081 (1982).

36. Laughlin, C. A., Tratschin, J. D., Coon, H. & Carter, B. J. Cloning of infectious adeno-associated virus genomes in bacterial plasmids. *Gene* **23**, 65–73 (1983).

37. 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).

38. Tratschin, J. D., West, M. H., Sandbank, T. & Carter, B. J. A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. *Mol. Cell. Biol.* **4**, 2072–2081 (1984).

39. Samulski, R. J. & Muzyczka, N. AAV-mediated gene therapy for research and therapeutic purposes. *Annu. Rev. Virol.* **1**, 427–451 (2014).

40. Li, C. & Samulski, R. J. Engineering adeno-associated virus vectors for gene therapy. *Nat. Rev. Genet.* **21**, 255–272 (2020).

41. Smith, R. H. Adeno-associated virus integration: Virus versus vector. *Gene Ther.* **15**, 817–822 (2008).

42. Srivastava, A., Lusby, E. W. & Berns, K. I. Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J. Virol.* **45**, 555–564 (1983).

43. 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).

44. Samulski, R. J., Srivastava, A., Berns, K. I. & Muzyczka, N. Rescue of adeno-associated virus from recombinant plasmids: Gene correction within the terminal repeats of AAV. *Cell* **33**, 135–143 (1983).

45. Yang, Q., Kadam, A. & Trempe, J. P. Mutational analysis of the adeno-associated virus rep gene. *J. Virol.* **66**, 6058–6069 (1992).

46. Hermonat, P. L., Labow, M. A., Wright, R., Berns, K. I. & Muzyczka, N. Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. *J. Virol.* **51**, 329–339 (1984).

47. Tratschin, J. D., Miller, I. L. & Carter, B. J. Genetic analysis of adeno-associated virus: properties of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function. *J. Virol.* **51**, 611–619 (1984).

48. Im, D. S. & Muzyczka, N. The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. *Cell* **61**, 447–457 (1990).

49. Mietzsch, M. *et al.* Improved Genome Packaging Efficiency of Adeno-associated Virus Vectors Using Rep Hybrids structure and assembly. *J. Virol.* **95**, 773–794 (2021).

50. Hong, G., Ward, P. & Berns, K. I. In vitro replication of adeno-associated virus DNA. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4673–4677 (1992).

51. 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).

52. 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).

53. 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).

54. 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).

55. 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).

56. Chejanovsky, N. & Carter, B. J. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: Effects on viral DNA replication. *Virology* **173**, 120–128 (1989).

57. Saudan, P. Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb. *EMBO J.* **19**, 4351–4361 (2000).

58. 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).

59. Berthet, C., Raj, K., Saudan, P. & Beard, P. How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13634–13639 (2005).

60. 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).
61. 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).
62. 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).
63. 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. *J. Virol.* **78**, 6595–6609 (2004).
64. Gonçalves, M. A. F. V. Adeno-associated virus: From defective virus to effective vector. *Virol. J.* **2**, 1–17 (2005).
65. Flotte, T. R. & Berns, K. I. Adeno-associated virus: A ubiquitous commensal of mammals. *Hum. Gene Ther.* **16**, 401–407 (2005).
66. Daya, S. & Berns, K. I. Gene therapy using adeno-associated virus vectors. *Clin. Microbiol. Rev.* **21**, 583–593 (2008).
67. Xie, Q. *et al.* The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 10405–10410 (2002).
68. Wu, Z., Asokan, A. & Samulski, R. J. Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy. *Mol. Ther.* **14**, 316–327 (2006).
69. Agbandje-McKenna, M. & Kleinschmidt, J. AAV Capsid Structure and Cell Interactions. *Adeno-Associated Virus: Methods and Protocols* **807**, 47-92 (2011).
70. Gurda, B. L. *et al.* Capsid Antibodies to Different Adeno-Associated Virus Serotypes Bind Common Regions. *J. Virol.* **87**, 9111–9124 (2013).
71. Lin-Ya Huang, Sujata Halder, M. A.-M. Parvovirus Glycan Interactions. *Curr. Opin. Virol.* **0**, 108–118 (2014).
72. Summerford, C. & Samulski, R. J. Membrane-Associated Heparan Sulfate Proteoglycan Is a Receptor for Adeno-Associated Virus Type 2 Virions. *J. Virol.* **72**, 1438–1445 (1998).
73. Verdera, H. C., Kuranda, K. & Mingozzi, F. AAV Vector Immunogenicity in Humans: A Long Journey to Successful Gene Transfer. *Mol. Ther.* **28**, 723–746 (2020).
74. Qing, K. *et al.* Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat. Med.* **5**, 71–77 (1999).
75. S. Pillay, NL Meyer, AS Puschnik, O Davulcu, J Diep, Y. I. An essential receptor for adeno-associated virus infection. *Nature* **530**, 139–148 (2016).

76. Dudek, A. M. *et al.* An Alternate Route for Adeno-associated Virus (AAV) Entry Independent of AAV Receptor. *J. Virol.* **92**, 1–15 (2018).

77. Erles, K., Sebökovà, P. & Schlehofer, J. R. Update on the prevalence of serum antibodies (IgG and IgM) to adeno- associated virus (AAV). *J. Med. Virol.* **59**, 406–411 (1999).

78. Monteilhet, V., Veron, P., Leborgne, C. & Benveniste, O. Prevalence of Serum IgG and Neutralizing Factors and 9 in the Healthy Population : Implications for Gene Therapy Using AAV Vectors. *Hum. Gene Ther.* **21**, 704–712 (2010).

79. Kruzik, A. *et al.* Prevalence of Anti-Adeno-Associated Virus Immune Responses in International Cohorts of Healthy Donors. *Mol. Ther. Methods Clin. Dev.* **14**, 126–133 (2019).

80. Calcedo, R. & Wilson, J. M. Humoral Immune Response to AAV. *Front. Immunol.* **4**, 1–7 (2013).

81. Patel, T. V. Process design for scalable recombinant adeno-associated virus. *Doctoral thesis, University of North Carolina* (2017).

82. 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).

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

84. 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).

85. 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**, 1980–1996 (2017).

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

87. 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).

88. 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. *Mol Bio.* **1143**, 1139–1143 (2019).

89. Elmore, Z. C. *et al.* The membrane associated accessory protein is an adeno-associated viral egress factor. *Nat. Commun.* **12**, 1–13 (2021).

90. Meier, A. F., Fraefel, C. & Seyffert, M. The interplay between adeno-associated virus and its helper viruses. *Viruses* **12**, 8–12 (2020).

91. 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).

92. 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).

93. Transgene, A. V. *et al.* Human Papillomavirus Type 16 Helper Functions for Adeno-Associated Virus Type 2 Replication. *Mol. Ther.* **9**, S289–S290 (2004).

94. Wang, Z. *et al.* Human Bocavirus 1 Is a Novel Helper for Adeno-associated Virus Replication. *J. Virol.* **91**, 1–18 (2017).

95. 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).

96. Yalkinoglu AO, Heilbronn R, Bürkle A, Schlehofer JR, zur H. H. DNA amplification of adeno-associated virus as a response to cellular genotoxic stress. *Cancer Res.* **48**, 3123–3129 (1988).

97. Richardson, W. D. & Westphal, H. A cascade of adenovirus early functions is required for expression of adeno-associated virus. *Cell.* **27**, 133–141 (1981).

98. Samulski, R. J., Chang, L. S. & Shenk, T. A recombinant plasmid from which an infectious adeno-associated virus genome can be excised in vitro and its use to study viral replication. *J. Virol.* **61**, 3096–3101 (1987).

99. 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**, 2224–2232 (1998).

100. Allen, J. M., Halbert, C. L. & Miller, A. D. Improved Adeno-Associated Virus Vector Production with Transfection of a Single Helper Adenovirus Gene, E4orf6. *Mol. Ther.* **1**, 88–95 (2000).

101. Matsushita, T. *et al.* The adenovirus E1A and E1B19K genes provide a helper function for transfection-based adeno-associated virus vector production. *J. Gen. Virol.* **85**, 2209–2214 (2004).

102. Bulcha, J. T., Wang, Y., Ma, H., Tai, P. W. L. & Gao, G. Viral vector platforms within the gene therapy landscape. *Signal Transduct. Target. Ther.* **6**, 1–24 (2021).

103. Richardsoni, W. D. & Westphal2, H. Requirement for Either Early Region Ia or Early Region Ib Adenovirus Gene Products in the Helper Effect for Adeno-Associated Virus. *J. Virol.* **51**,

404–410 (1984).

104. Chang, L. S., Shi, Y. & Shenk, T. Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. *J. Virol.* **63**, 3479–3488 (1989).
105. Mouw, M. B. & Pintel, D. J. Adeno-Associated Virus RNAs Appear in a Temporal Order and Their Splicing Is Stimulated during Coinfection with Adenovirus. *J. Virol.* **74**, 9878–9888 (2000).
106. Matsushita, T. *et al.* Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Ther.* **5**, 938–945 (1998).
107. Lillie, J. W., Loewenstein, P. M., Green, M. R. & Green, M. Functional domains of adenovirus type 5 E1a proteins. *Cell.* **50**, 1091–1100 (1987).
108. Pelka, P. *et al.* Transcriptional control by adenovirus E1A conserved region 3 via p300/CBP. *Nucleic Acids Res.* **37**, 1095–1106 (2009).
109. Berk, A. Adenovirus Promoters and E1A Transactivation. *Annu. Rev. Genet.* **20**, 45–79 (1986).
110. McConnell, M. J. & Imperiale, M. J. Biology of adenovirus and its use as a vector for gene therapy. *Hum. Gene Ther.* **15**, 1022–1033 (2004).
111. Pötzer, B. M. *et al.* E1A is sufficient by itself to induce apoptosis independent of p53 and other adenoviral gene products. *Cell Death Differ.* **7**, 177–188 (2000).
112. White, E. Function of the adenovirus E1B oncogene in infected and transformed cells. *Semin. Virol.* **5**, 341–348 (1994).
113. Martlnout, I., Raven, T., Pun, K. & Grinham, C. J. Cloning of a bcl-2 homologue by interaction with adenovirus E1B 19K. *Nature.* **375**, 731–733 (1995).
114. White, E. Regulation of the cell cycle and apoptosis by the oncogenes of adenovirus. *Oncogene.* **20**, 7836–7846 (2001).
115. Rao, L. *et al.* The adenovirus E1A proteins induce apoptosis, which is inhibited by the E1B 19-kDa and Bcl-2 proteins. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7742–7746 (1992).
116. Leppard, K. N. E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. *J. Gen. Virol.* **78**, 2131–2138 (1997).
117. Pilder, S., Moore, M., Logan, J. & Shenk, T. The Adenovirus E1B-55K Transforming Polypeptide Modulates Transport or cytoplasmic stabilization of viral and host cell mRNAs. *Mol. Cell. Biol.* **6**, 470–476 (1986).
118. Millet, R. *et al.* Impact of the MRN Complex on Adeno-Associated Virus Integration and Replication during Coinfection with Herpes Simplex Virus 1. *J. Virol.* **89**, 6824–6834 (2015).

119. Schwartz, R. A. *et al.* The Mre11/Rad50/Nbs1 complex limits adeno-associated virus transduction and replication. *J. Virol.* **81**, 12936–12945 (2007).

120. Mohammadi, E. S., Ketner, E. A., Johns, D. C. & Ketner, G. Expression of the adenovirus E4 34k oncoprotein inhibits repair of double strand breaks in the cellular genome of a 293-based inducible cell line. *Nucleic Acids Res.* **32**, 2652–2659 (2004).

121. Chang, L. S. & 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).

122. Hidalgo, P. *et al.* Morphological, Biochemical, and Functional Study of Viral Replication Compartments Isolated from Adenovirus-Infected Cells. *J. Virol.* **90**, 3411–3427 (2016).

123. Ward, P., Dean, F. B., O'Donnell, M. E. & Berns, K. I. Role of the Adenovirus DNA-Binding Protein in In Vitro Adeno-Associated Virus DNA Replication. *J. Virol.* **72**, 420–427 (1998).

124. Weitzman, M. D., Fisher, K. J. & Wilson, J. M. Recruitment of wild-type and recombinant adeno-associated virus into adenovirus replication centers. *J. Virol.* **70**, 1845–1854 (1996).

125. Swaminathan, S. & Thimmapaya, B. Transactivation of adenovirus E2-early promoter by E1A and E4 6/7 in the context of viral chromosome. *J. Mol. Biol.* **258**, 736–746 (1996).

126. Klessig, D. F., Brough, D. E. & Cleghon, V. Introduction, Stable Integration, and Controlled Expression of a Chimeric Adenovirus Gene Whose Product Is Toxic to the Recipient Human Cell. *Mol. Cell. Biol.* **4**, 1354–1362 (1984).

127. Zhou, H., O'Neal, W., Morral, N. & Beaudet, A. L. Development of a complementing cell line and a system for construction of adenovirus vectors with E1 and E2a deleted. *J. Virol.* **70**, 7030–7038 (1996).

128. Marcellus, R. C. *et al.* The Early Region 4 orf4 Protein of Human Adenovirus Type 5 Induces p53-Independent Cell Death by Apoptosis. *J. Virol.* **72**, 7144–7153 (1998).

129. Querido, E. *et al.* Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. *Genes Dev.* **15**, 3104–3117 (2001).

130. Samulski, R. J. & Shenk, T. Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *J. Virol.* **62**, 206–210 (1988).

131. Shi, Y., Seto, E., Chang, L. S. & Shenk, T. Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. *Cell.* **67**, 377–388 (1991).

132. Kitajewski, J. *et al.* Adenovirus VAI RNA antagonizes the antiviral action of interferon by preventing activation of the interferon-induced eIF-2α kinase. *Cell.* **45**, 195–200 (1986).

133. Vachon, V. K. & Conn, G. L. Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Res.* **212**, 39–52 (2016).

134. O'Malley, R. P., Mariano, T. M., Siekierka, J. & Mathews, M. B. A mechanism for the control of protein synthesis by adenovirus VA RNA I. *Cell.* **44**, 391–400 (1986).

135. Fu, Y. *et al.* Comprehensive biophysical characterization of AAV-AAVR interaction uncovers serotype- and pH-dependent interaction. *J. Pharm. Biomed. Anal.* **234**, 1–8 (2023).

136. Bartlett, J. S., Wilcher, R. & R. Jude. Infectious Entry Pathway of Adeno-Associated Virus and Adeno-Associated Virus Vectors. *J. Virol.* **74**, 2777–2785 (2000).

137. Duan, D. *et al.* Dynamin Is Required for Recombinant Adeno-Associated Virus Type 2 Infection. *J. Virol.* **73**, 10371–10376 (1999).

138. Nonnenmacher, M. & Weber, T. Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. *Cell Host Microbe.* **10**, 563–576 (2011).

139. Bantel-Schaal, U., Hub, B. & Kartenbeck, J. Endocytosis of Adeno-Associated Virus Type 5 Leads to Accumulation of Virus Particles in the Golgi Compartment. *J. Virol.* **76**, 2340–2349 (2002).

140. Bartlett, J. S., Wilcher, R., R. Jude & Samulski, R. J. Infectious Entry Pathway of Adeno-Associated Virus and Adeno-Associated Virus Vectors. *J. Virol.* **74**, 2777–2785 (2000).

141. 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**, 10462–10473 (2012).

142. 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).

143. Kronenberg, S., Bottcher, B., von der Lieth, C. W., 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).

144. Johnson, J. S. & Samulski, R. J. Enhancement of Adeno-Associated Virus Infection by Mobilizing Capsids into and Out of the Nucleolus. *J. Virol.* **83**, 2632–2644 (2009).

145. Nicolson, S. C. & Samulski, R. J. Recombinant Adeno-Associated Virus Utilizes Host Cell Nuclear Import Machinery To Enter the Nucleus. *J. Virol.* **88**, 4132–4144 (2014).

146. Sutter, S. O. *et al.* Adeno-associated virus type 2 (AAV2) uncoating is a stepwise process and is linked to structural reorganization of the nucleolus. *PLoS Pathog.* **18**, 1–30 (2022).

147. Nash, K., Chen, W. & Muzyczka, N. Complete In Vitro Reconstitution of Adeno-Associated

Virus DNA Replication Requires the Minichromosome Maintenance Complex Proteins. *J. Virol.* **82**, 1458–1464 (2008).

148. Charman, M. & Weitzman, M. D. Replication compartments of DNA viruses in the nucleus: Location, location, location. *Viruses.* **12**, 1–20 (2020).
149. Bennett, A., Mietzsch, M. & Agbandje-Mckenna, M. Understanding capsid assembly and genome packaging for adeno-associated viruses. *Futur. Virol.* **12**, 283–297 (2017).
150. Smith, G. A. & Enquist, L. W. Break ins and break outs: Viral interactions with the cytoskeleton of mammalian cells. *Annu. Rev. Cell Dev. Biol.* **18**, 135–161 (2002).
151. Lock, M. *et al.* Rapid, simple, and versatile manufacturing of recombinant adeno-associated viral vectors at scale. *Hum. Gene Ther.* **21**, 1259–1271 (2010).
152. Vandenberghe, L. H. *et al.* Efficient serotype-dependent release of functional vector into the culture medium during adeno-associated virus manufacturing. *Hum. Gene Ther.* **21**, 1251–1257 (2010).
153. Schiller, L. T., Lemus-Diaz, N., Rinaldi Ferreira, R., Böker, K. O. & Gruber, J. Enhanced Production of Exosome-Associated AAV by Overexpression of the Tetraspanin CD9. *Mol. Ther. - Methods Clin. Dev.* **9**, 278–287 (2018).
154. György, B. & Maguire, C. A. Extracellular vesicles: nature's nanoparticles for improving gene transfer with adeno-associated virus vectors. *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology.* **10**, 1–13 (2018).
155. Hudry, E. *et al.* Exosome-associated AAV vector as a robust and convenient neuroscience tool. *Gene Ther.* **23**, 380–392 (2016).
156. Xiao, X. Adeno-associated viral vectors found free in media. *Hum. Gene Ther.* **21**, 1221–1222 (2010).
157. Smith-Moore, S. *et al.* Adeno-associated virus Rep proteins antagonize phosphatase PP1 to counteract KAP1 repression of the latent viral genome. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3529–E3538 (2018).
158. Lee, J. S. *et al.* Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes Dev.* **9**, 1188–1198 (1995).
159. Schnepp, B. C., Jensen, R. L., Clark, K. R. & Johnson, P. R. Infectious Molecular Clones of Adeno-Associated Virus Isolated Directly from Human Tissues. *J. Virol.* **83**, 1456–1464 (2009).
160. Duan, D. *et al.* Circular Intermediates of Recombinant Adeno-Associated Virus Have Defined Structural Characteristics Responsible for Long-Term Episomal Persistence in Muscle Tissue. *J. Virol.* **73**, 861–861 (1999).

161. Duan, D., Yan, Z., Yue, Y. & Engelhardt, J. F. Structural analysis of adeno-associated virus transduction circular intermediates. *Virology.* **261**, 8–14 (1999).

162. Tyson, J. J., Chen, K. C., Lederman, M. & Bates, R. C. Analysis of the kinetic hairpin transfer model for parvoviral DNA replication. *J. Theor. Biol.* **144**, 155–169 (1990).

163. Samulski, R. J. *et al.* Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J.* **10**, 3941–3950 (1991).

164. Kotin, R. M., Linden, R. M. & Berns, K. I. 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).

165. Linden, R. M., Winocour, E. & Berns, K. I. The recombination signals for adeno-associated virus site-specific integration. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 7966–7972 (1996).

166. 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).

167. Meneses, P., Berns, K. I. & Winocour, E. DNA Sequence Motifs Which Direct Adeno-Associated Virus Site-Specific Integration in a Model System. *J. Virol.* **74**, 6213–6216 (2000).

168. Janovitz, T. *et al.* High-Throughput Sequencing Reveals Principles of Adeno-Associated Virus Serotype 2 Integration. *J. Virol.* **87**, 8559–8568 (2013).

169. Vincent-Lacaze, N. *et al.* Structure of Adeno-Associated Virus Vector DNA following Transduction of the Skeletal Muscle. *J. Virol.* **73**, 1949–1955 (1999).

170. Penaud-Budloo, M. *et al.* Adeno-Associated Virus Vector Genomes Persist as Episomal Chromatin in Primate Muscle. *J. Virol.* **82**, 7875–7885 (2008).

171. Nowrouzi, A. *et al.* Integration frequency and intermolecular recombination of rAAV vectors in non-human primate skeletal muscle and liver. *Mol. Ther.* **20**, 1177–1186 (2012).

172. Nakai, H. *et al.* AAV serotype 2 vectors preferentially integrate into active genes in mice. *Nat. Genet.* **34**, 297–302 (2003).

173. Chandler, R. J., LaFave, M. C., Varshney, G. K., Burgess, S. M. & Venditti, C. P. Genotoxicity in mice following AAV gene delivery: A safety concern for human gene therapy? *Mol. Ther.* **24**, 198–201 (2016).

174. McCarty, D. M., Young, S. M. & Samulski, R. J. Integration of Adeno-Associated Virus (AAV) and Recombinant AAV Vectors. *Annu. Rev. Genet.* **38**, 819–845 (2004).

175. Dong, J., Fan, P., Frizzell, R. A. Quantitative analysis of the packaging capacity of recombinant adeno-associated virus. *Hum Gen Ther.* **2112**, 2101–2112 (1996).

176. Wang, B., Li, J. & Xiao, X. Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13714–13719 (2000).

177. Wang, D., Tai, P. W. L. & Gao, G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* **18**, 358–378 (2019).

178. Grimm, D., Pandey, K., Nakai, H., Storm, T. A. & Kay, M. A. Liver Transduction with Recombinant Adeno-Associated Virus Is Primarily Restricted by Capsid Serotype Not Vector Genotype. *J. Virol.* **80**, 426–439 (2006).

179. Samulski, R. J., Chang, L. S. & Shenk, T. Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *J. Virol.* **63**, 3822–3828 (1989).

180. Allen, J. M., Debelak, D. J., Reynolds, T. C. & Miller, A. D. Identification and elimination of replication-competent adeno-associated virus (AAV) that can arise by nonhomologous recombination during AAV vector production. *J. Virol.* **71**, 6816–6822 (1997).

181. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36**, 59–72 (1977).

182. 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).

183. European Medicines Agency. 'Zynteglo: Withdrawal of the marketing authorisation in the European Union' EMA/192892/2022. (2022). Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/zynteglo. (Accessed: 20th February 2024)

184. Zolgensma® (onasemnogene abeparvovec) prescribing information. (2019). Available at: https://www.fda.gov/vaccines-blood-biologics/zolgensma. (Accessed: 1st September 2023)

185. Hemgenix® (etranacogene dezaparvovec) prescribing information. (2022). Available at: https://labeling.cslbehring.com/PI/US/Hemgenix/EN/Hemgenix-Prescribing-Information.pdf. (Accessed: 1st September 2023)

186. 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).

187. Sha, S. *et al.* Cellular pathways of recombinant adeno-associated virus production for gene therapy. *Biotechnol. Adv.* **49**, 107764 (2021).

188. Ayuso, E., Mingozzi, F. & Bosch, F. Production, Purification and Characterization of Adeno-Associated Vectors. *Curr. Gene Ther.* **10**, 423–436 (2010).

189. 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).

190. Gutiérrez-Granados, S., Cervera, L., Kamen, A. A. & Gòdia, F. Advancements in mammalian cell transient gene expression (TGE) technology for accelerated production of biologics. *Crit. Rev. Biotechnol.* **38**, 918–940 (2018).
191. Cervera, L., González-Domínguez, I., Segura, M. M. & Gòdia, F. Intracellular characterization of Gag VLP production by transient transfection of HEK 293 cells. *Biotechnol. Bioeng.* **114**, 2507–2517 (2017).
192. Riley, M. K. & Vermerris, W. Recent Advances in Nanomaterials for Gene Delivery—A Review. *Nanomaterials* **7**, 1–19 (2017).
193. Geisse, S. & Fux, C. Chapter 15 Recombinant Protein Production by Transient Gene Transfer into Mammalian Cells. *Methods Enzymol.* **463**, 223–238 (2009).
194. Srivastava, A., Mallela, K. M. G., Deorkar, N. & Brophy, G. Manufacturing challenges and rational formulation development for AAV viral vectors. *J. Pharm. Sci.* **110**, 2609–2624 (2021).
195. 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).
196. Urabe, M. *et al.* Scalable Generation of High-Titer Recombinant Adeno-Associated Virus Type 5 in Insect Cells. *J. Virol.* **80**, 1874–1885 (2006).
197. Yang, Q., Chen, F. & Trempe, J. P. Characterization of cell lines that inducibly express the adeno- associated virus rep proteins. *J. Virol.* **68**, 4847–4856 (1994).
198. Marcellus, R. C. *et al.* Adenovirus type 5 early region 4 is responsible for E1A-induced p53-independent apoptosis. *J. Virol.* **70**, 6207–6215 (1996).
199. Luxturna® (vortigene neparvovec-rzyl) prescribing information. (2017). Available at: https://www.fda.gov/media/109906/download. (Accessed: 1st February 2023)
200. *European Medicines Agency. 'Public Assessment Report (EPAR): Zolgensma.' EMA/200482/2020*. (2022).
201. *BLA - FDA Advisory commitee briefing document: LUXTURNA TM (voretigene neparvovec)*. *FDA* 1–88 (2017).
202. FDA Advisory Committee. 'Briefing Document: Luxturna™ (vortigene neparvovec).' Spark Therapeutics, Inc. 1–149 (2017).
203. Powers, A. D., Piras, B. A., Clark, R. K., Lockey, T. D. & Meagher, M. M. Development and Optimization of AAV hFIX Particles by Transient Transfection in an iCELLis ® Fixed-Bed

Bioreactor. *Hum. Gene Ther. Methods.* **27**, 112–121 (2016).

204. 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).
205. Côté, J., Garnier, A., Massie, B., Kamen, A. Serum-free production of recombinant proteins and adenoviral vectors by 293SF-3F6 cells. *Biotechnol. Bioeng.* **59**, 567–575 (1998).
206. Thermo Fisher Scientific. Expi293™ Expression System for Transient Protein Expression. Available at: https://www.thermofisher.com/es/es/home/life-science/protein-biology/protein-expression/mammalian-protein-expression/transient-mammalian-protein-expression/expi293-expression-system.html#capabilities. (Accessed: 1st April 2024)
207. 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* **196**, 163–173 (2014).
208. Durocher, Y. *et al.* Scalable serum-free production of recombinant adeno-associated virus type 2 by transfection of 293 suspension cells. *J. Virol. Methods* **144**, 32–40 (2007).
209. Joon Young Park, Byung-Pil Lim, Kyuhyun Lee, Young-Gun Kim, E.-C. J. Scalable Production of Adeno-Associated Virus Type 2 Vectors Via Suspension Transfection. *Biotechnol. Bioeng.* **94**, 416–430 (2006).
210. 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).
211. van Lieshout, L. P. *et al.* A novel dual-plasmid platform provides scalable transfection yielding improved productivity and packaging across multiple AAV serotypes and genomes. *Mol. Ther. Methods Clin. Dev.* **29**, 426–436 (2023).
212. Zhao, H. *et al.* Creation of a High-Yield AAV Vector Production Platform in Suspension Cells Using a Design-of-Experiment Approach. *Mol. Ther. - Methods Clin. Dev.* **18**, 312–320 (2020).
213. Collaud, F. *et al.* Preclinical Development of an AAV8-hUGT1A1 Vector for the Treatment of Crigler-Najjar Syndrome. *Mol. Ther. Methods Clin. Dev.* **12**, 157–174 (2019).
214. Godia-Casablancas, Francesc; Bosch-Tubert, Maria-Fatima; Garcia-Martinez, Miguel; Cervera-Gracia, Laura; Leon-Madrenas, Xavier; Molas-Laplana, Maria; Gutierrez-Granados, S. Methods for the manufacture of recombinant viral vectors (EP2020058527). European Patent Office. (2020).
215. Brimble, M. A. *et al.* Preventing packaging of translatable P5-associated DNA contaminants

in recombinant AAV vector preps. *Mol. Ther. Methods Clin. Dev.* **24**, 280–291 (2022).

216. Grimm, D., Kern, A., Rittner, K. & Kleinschmidt, J. A. Novel tools for production and purification of recombinant adenoassociated virus vectors. *Hum. Gene Ther.* **9**, 2745–2760 (1998).
217. Tang, Q. *et al.* Two-plasmid packaging system for recombinant adeno-associated virus. *Biores. Open Access* **9**, 219–228 (2020).
218. Mauro, E. *et al.* pPLUS ® AAV-Helper, novel engineered pHelper plasmid to improve yield and quality of several AAV serotypes in suspension cell culture systems. *Poster by Polyplus*
219. Wang, X.-S. *et al.* Characterization of Wild-Type Adeno-Associated Virus Type 2-Like Particles Generated during Recombinant Viral Vector Production and Strategies for Their Elimination. *J. Virol.* **72**, 5472–5480 (1998).
220. Wright, J. F. Product-related impurities in clinical-grade recombinant AAV vectors: Characterization and risk assessment. *Biomedicines* **2**, 80–97 (2014).
221. Cao, L., Liu, Y., During, M. J. & Xiao, W. High-Titer, Wild-Type Free Recombinant Adeno-Associated Virus Vector Production Using Intron-Containing Helper Plasmids. *J. Virol.* **74**, 11456–11463 (2000).
222. Conway, J. E. *et al.* High-titer recombinant adeno-associated virus production utilizing a recombinant herpes simplex virus type I vector expressing AAV-2 Rep and Cap. *Gene Ther. 1999 66* **6**, 986–993 (1999).
223. Thomas, D. L. *et al.* Scalable Recombinant Adeno-Associated Virus Production Using Recombinant Herpes Simplex Virus Type 1 Coinfection of Suspension-Adapted Mammalian Cells. *Hum. Gene Ther.* **20**, 861–870 (2009).
224. Stutika, C. *et al.* Definition of herpes simplex virus helper functions for the replication of adeno-associated virus type 5. *J. Gen. Virol.* **96**, 840–850 (2015).
225. Knop, D. R. & Harrell, H. Bioreactor production of recombinant herpes simplex virus vectors. *Biotechnol. Prog.* **23**, 715–721 (2007).
226. Knop DR, Thomas DL, B. C. 593. SFM rHSV-based production of rAAV vectors in suspension BHK cells. *Mol Ther* **19**, (2011).
227. Aponte-Ubillus, J. J. *et al.* Molecular design for recombinant adeno-associated virus (rAAV) vector production. *Appl. Microbiol. Biotechnol.* **102**, 1045–1054 (2018).
228. Kost, T. A., Patrick Condreay, J. & Jarvis, D. L. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat. Biotechnol. Vol.* **23**, 567-575 (2005).
229. Jeang, K. T., Holmgren-Konig, M. & Khoury, G. A baculovirus vector can express intron-containing genes. *J. Virol.* **61**, 1761–1764 (1987).

230. *European Medicines Agency. 'Public Assessment Report (EPAR): Hemgenix.' EMA/46569/2023.* (2023).

231. *European Medicines Agency. 'Public Assessment Report (EPAR): Roctavian.' EMA/685615/2022*. (2022).

232. Kohlbrenner, E. *et al.* Successful Production of Pseudotyped rAAV Vectors Using a Modified Baculovirus Expression System. *Mol. Ther.* **12**, 1217–1225 (2005).

233. Chen, H. Intron splicing-mediated expression of AAV rep and cap genes and production of AAV vectors in insect cells. *Mol. Ther.* **16**, 924–930 (2008).

234. Smith, R. H., Levy, J. R. & Kotin, R. M. A Simplified Baculovirus-AAV Expression Vector System Coupled With One-step Affinity Purification Yields High-titer rAAV Stocks From Insect Cells. *Am. Soc. Gene Cell Ther.* **17**, 1888–1896 (2009).

235. Bosma, B. *et al.* Optimization of viral protein ratios for production of rAAV serotype 5 in the baculovirus system. *Gene Ther.* **25**, 415–424 (2018).

236. 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).

237. Joshi, P. R. H., Venereo-Sanchez, A., Chahal, P. S. & Kamen, A. A. Advancements in molecular design and bioprocessing of recombinant adeno-associated virus gene delivery vectors using the insect-cell baculovirus expression platform. *Biotechnol. J.* **16**, 1–17 (2021).

238. Cecchini, S., Virag, T. & Kotin, R. M. Reproducible high yields of recombinant adeno-associated virus produced using invertebrate cells in 0.02-to 200-liter cultures. *Hum. Gene Ther.* **22**, 1021–1030 (2011).

239. Luo, W.-Y. *et al.* Adaptive Immune Responses Elicited by Baculovirus and Impacts on Subsequent Transgene Expression In Vivo . *J. Virol.* **87**, 4965–4973 (2013).

240. Rumachik, N. G. *et al.* Methods Matter: Standard Production Platforms for Recombinant AAV Produce Chemically and Functionally Distinct Vectors. *Mol. Ther. - Methods Clin. Dev.* **18**, 98–118 (2020).

241. Aslanidi, G., Lamb, K. & Zolotukhin, S. An inducible system for highly efficient production of recombinant adeno-associated virus (rAAV) vectors in insect Sf9 cells. *PNAS* **106**, 5059–5064 (2009).

242. Rao, R. *et al.* Performance of Cardiotropic rAAV Vectors Is Dependent on Production Method. *Viruses* **14**, 1–12 (2022).

243. Van Ree, R. Carbohydrate epitopes and their relevance for the diagnosis and treatment of

allergic diseases. *Int. Arch. Allergy Immunol.* **129**, 189–197 (2002).

244. Tran, N. T. *et al.* Human and Insect Cell-Produced Recombinant Adeno-Associated Viruses Show Differences in Genome Heterogeneity. *Hum. Gene Ther.* **33**, 371–388 (2022).
245. Gu, B. *et al.* Establishment of a scalable manufacturing platform for in-silico-derived ancestral adeno-associated virus. *Cell Gene Ther. Insigths.* 753–769 (2018).
246. Fu, Q., Polanco, A., Lee, Y. S. & Yoon, S. Critical challenges and advances in recombinant adeno-associated virus (rAAV) biomanufacturing. *Biotechnol. Bioeng.* **120**, 2601–2621 (2023).
247. Jalšić, L. *et al.* Inducible HEK293 AAV packaging cell lines expressing Rep proteins. *Mol. Ther. - Methods Clin. Dev.* **30**, 259–275 (2023).
248. 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).
249. 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).
250. Inoue, N. & Russell, D. W. Packaging Cells Based on Inducible Gene Amplification for the Production of Adeno-Associated Virus Vectors. *J. Virol.* **72**, 7024–7031 (1998).
251. Gao, G. *et al.* High-Titer Adeno-Associated Viral Vectors from Shuttle Virus a Rep/Cap Cell line and Hybrid. **2362**, 2353–2362 (1998).
252. Chadeuf, G. *et al.* Efficient recombinant adeno-associated virus production by a stable rep-cap HeLa cell line correlates with adenovirus-induced amplification of the integrated rep-cap genome. *J. Gene Med.* **2**, 260–268 (2000).
253. Clark, K. R., Voulgaropoulou, F., Fraley, D. M., Johnsoni, P. R. & Al, C. E. T. Cell lines for the production of recombinant adeno-associated virus. *Hum Gene Ther* **1341**, 1329–1341 (1995).
254. Gao, G. P. *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).
255. Ogasawara, Y. *et al.* Highly regulated expression of adeno-associated virus large Rep proteins in stable 293 cell lines using the Cre/loxP switching system. *J. Gen. Virol.* **80**, 2477–2480 (1999).
256. Qiao, C., Wang, B., Zhu, X., Li, J. & Xiao, X. A Novel Gene Expression Control System and Its Use in Stable, High-Titer 293 Cell-Based Adeno-Associated Virus Packaging Cell Lines. *J. Virol.* **76**, 13015–13027 (2002).

257. Qiao, C., Li, J., Skold, A., Zhang, X. & Xiao, X. Feasibility of Generating Adeno-Associated Virus Packaging Cell Lines Containing Inducible Adenovirus Helper Genes. *J. Virol.* **76**, 1904–1913 (2002).

258. Hölscher, C. *et al.* Cell lines inducibly expressing the adeno-associated virus (AAV) rep gene: Requirements for productive replication of rep-negative AAV mutants. *J. Virol.* **68**, 7169–7177 (1994).

259. Okada, T. *et al.* Development and characterization of an antisense-mediated prepackaging cell line for adeno-associated virus vector production. *Biochem. Biophys. Res. Commun.* **288**, 62–68 (2001).

260. 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).

261. Liu, X., Voulgaropoulou, F., Chen, R., Johnson, P. R. & Clark, K. R. Selective Rep-Cap gene amplification as a mechanism for high-titer recombinant AAV production from stable cell lines. *Mol. Ther.* **2**, 394–403 (2000).

262. Nony, P. *et al.* Novel cis -Acting Replication Element in the Adeno-Associated Virus Type 2 Genome Is Involved in Amplification of Integrated rep-cap Sequences . *J. Virol.* **75**, 9991–9994 (2001).

263. Tessier, J. *et al.* Characterization of Adenovirus-Induced Inverted Terminal Repeat-Independent Amplification of Integrated Adeno-Associated Virus rep-cap Sequences . *J. Virol.* **75**, 375–383 (2001).

264. Su, W., Seymour, L. W. & Cawood, R. AAV production in stable packaging cells requires expression of adenovirus 22/33K protein to allow episomal amplification of integrated rep/cap genes. *Sci. Rep.* **13**, 1–13 (2023).

265. Thorne, B. A., Takeya, R. K. & Peluso, R. W. Manufacturing recombinant adeno-associated viral Vectors from Producer Cell Clones. **714**, 707–714 (2009).

266. Liu, X. L., Clark, K. R. & Johnson, P. R. Production of recombinant adeno-associated virus vectors using a packaging cell line and a hybrid recombinant adenovirus. *Gene Ther.* **6**, 293–299 (1998).

267. Escandell, J. *et al.* Towards a scalable bioprocess for rAAV production using a HeLa stable cell line. *Biotechnol. Bioeng.* **120**, 2578–2587 (2023).

268. U.S. Department of Health and Human Services, Food and Drug Administration & Center for Biologics Evaluation and Research. 'Guidance for Industry: Preclinical Assessment of Investigational Cellular and Gene Therapy Products'. (2024).

269. Mietzsch, M., Casteleyn, V., Weger, S., Zolotukhin, S. & Heilbronn, R. OneBac 2.0: Sf9 Cell Lines for Production of AAV5 Vectors with Enhanced Infectivity and Minimal Encapsidation of Foreign DNA. *Hum. Gene Ther.* **26**, 688 (2015).

270. 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).

271. 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).

272. Joshi, P. R. H. *et al.* Achieving High-Yield Production of Functional AAV5 Gene Delivery Vectors via Fedbatch in an Insect Cell-One Baculovirus System. *Mol. Ther. Methods Clin. Dev.* **13**, 279–289 (2019).

273. Wu, Y. *et al.* Development of versatile and flexible sf9 packaging cell line-dependent onebac system for large-scale recombinant adeno-associated virus production. *Hum. Gene Ther. Methods* **30**, 172–183 (2019).

274. Hein, K., Wissing, S., Riebesehl, N., Strempel, N., Faust, N., & Graßl, M. Generation of helper virus-free adeno-associated viral vector packaging/producer cell lines based on a human suspension cell line. *Proc. Cell Eng. Conf. XVI* Poster number 136 (2018).

275. Lee, Z., Lu, M., Irfanullah, E., Soukup, M. & Hu, W. S. Construction of an rAAV Producer Cell Line through Synthetic Biology. *ACS Synth. Biol.* **11**, 3285–3295 (2022).

276. Smith, R. H., Ding, C. & Kotin, R. M. Serum-free production and column purification of adeno-associated virus type 5. *J. Virol. Methods* **114**, 115–124 (2003).

277. Chung, C. H. *et al.* Production of rAAV by plasmid transfection induces antiviral and inflammatory responses in suspension HEK293 cells. *Mol. Ther. - Methods Clin. Dev.* **28**, 272–283 (2023).

278. Todd P Sanderson, Timothy Erlandson, Nathan Hazi, Anne MacIntyre, Benjamin I Ingersoll, Michael McLaughlin, Steven Wesel, P. B. M. *et al.* Scalability comparison between 50 and 500 liter stirred tank bioreactor for production of rAAV viral vecto. *Cell Gene Ther. Insights* **7**, 1025–1033 (2021).

279. Wright, J. F. Transient transfection methods for clinical adeno-associated viral vector production. *Hum. Gene Ther.* **20**, 698–706 (2009).

280. van der Loo, J. C. M. & Wright, J. F. Progress and challenges in viral vector manufacturing. *Hum. Mol. Genet.* **25**, R42–R52 (2016).

281. Lavado-García, J., Pérez-Rubio, P., Cervera, L. & Gòdia, F. The cell density effect in animal cell-based bioprocessing: Questions, insights and perspectives. *Biotechnol. Adv.* **60**, 1–14

(2022).

282. Petiot, E., Cuperlovic-Culf, M., Shen, C. F. & Kamen, A. Influence of HEK293 metabolism on the production of viral vectors and vaccine. *Vaccine* **33**, 5974–5981 (2015).
283. Yang, S., Zhou, X., Li, R., Fu, X. & Sun, P. Optimized PEI-based Transfection Method for Transient Transfection and Lentiviral Production. *Curr. Protoc. Chem. Biol.* **9**, 147–157 (2017).
284. Robert, M. A. *et al.* Manufacturing of recombinant adeno-associated viruses using mammalian expression platforms. *Biotechnol. J.* **12**, 1–16 (2017).
285. Henry, O., Perrier, M. & Kamen, A. Metabolic flux analysis of HEK-293 cells in perfusion cultures for the production of adenoviral vectors. *Metab. Eng.* **7**, 467–476 (2005).
286. Mendes, J. P. *et al.* AAV process intensification by perfusion bioreaction and integrated clarification. *Front. Bioeng. Biotechnol.* **10**, 1–11 (2022).
287. Cervera, L. *et al.* Generation of HIV-1 Gag VLPs by transient transfection of HEK 293 suspension cell cultures using an optimized animal-derived component free medium. *J. Biotechnol.* **166**, 152–165 (2013).
288. Liu, Y. *et al.* Using a fed-batch culture strategy to enhance rAAV production in the baculovirus / insect cell system. *JBIOSC* **110**, 187–193 (2010).
289. Taghavian, O. *et al.* Development of a Feed Medium to Improve AAV Production in HEK293 Fed-Batch Culture. Poster at ESACT by Fujifilm Irvine Scientific (2022).
290. Klimpel, M. *et al.* Development of a perfusion process for continuous lentivirus production using stable suspension producer cell lines. *Biotechnol. Bioeng.* **120**, 2622–2638 (2023).
291. Karim Ghani, Alain Garnier, Helene Coelho, Julia Transfiguracion, Pierre Trudel, A. K. Retroviral Vector Production Using Suspension-Adapted 293GPG Cells in a 3L Acoustic Filter-Based Perfusion Bioreactor. *Biotechnol. Bioeng.* **95**, 653–660 (2006).
292. Valérie Cortin, Jules Thibault, Danielle Jacob, and A. G. High-Titer Adenovirus Vector Production in 293S Cell Perfusion Culture. *Biotechnol. Progr.* **20**, 858–863 (2004).
293. Clément, N. & Grieger, J. C. Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Mol. Ther. - Methods Clin. Dev.* **3**, 16002 (2016).
294. Cervera, L., Gutiérrez-Granados, S., Berrow, N. S., Segura, M. M. & Gòdia, F. Extended gene expression by medium exchange and repeated transient transfection for recombinant protein production enhancement. *Biotechnol. Bioeng.* **112**, 934–946 (2015).
295. Fuenmayor, J., Cervera, L., Gòdia, F. & Kamen, A. Extended gene expression for Gag VLP production achieved at bioreactor scale. *J. Chem. Technol. Biotechnol.* **94**, 302–308 (2019).
296. Garnier, A., Côté, J., Nadeau, I., Kamen, A. & Massie, B. Scale-up of the adenovirus

expression system for the production of recombinant protein in human 293S cells. *Cytotechnology* **15**, 145–155 (1994).

297. Lavado-García, J., Cervera, L. & Gòdia, F. An Alternative Perfusion Approach for the Intensification of Virus-Like Particle Production in HEK293 Cultures. *Front. Bioeng. Biotechnol.* **8**, 1–16 (2020).
298. Bernal, V., Carinhas, N., Yokomizo, A. Y., Carrondo, M. J. T. & Alves, P. M. Cell density effect in the baculovirus-insect cells system: A quantitative analysis of energetic metabolism. *Biotechnol. Bioeng.* **104**, 162–180 (2009).
299. Wang, Y., Fu, Q., Lee, Y. S., Sha, S. & Yoon, S. Transcriptomic features reveal molecular signatures associated with recombinant adeno-associated virus production in HEK293 cells. *Biotechnol. Prog.* **e3346**, 1–13 (2023).
300. Strasser, L. *et al.* Proteomic landscape of adeno-associated virus (AAV)-producing HEK293 cells. *Int. J. Mol. Sci.* **22**, 1–14 (2021).
301. Lee, Z., Hu, W. S., Lu, M., Lee, Z. & Hu, W. S. Multi-omics kinetic analysis of recombinant adeno-associated virus production by plasmid transfection of HEK293 cells. *Biotechnol. Prog.* **e3428**, 1–12 (2024).
302. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).
303. Martínez-Bartolomé, S. *et al.* Properties of average score distributions of SEQUEST: the probability ratio method. *Mol. Cell. Proteomics* **7**, 1135–45 (2008).
304. Bonzon-Kulichenko, E., Garcia-Marques, F., Trevisan-Herraz, M. & Vazquez, J. Revisiting peptide identification by high-accuracy mass spectrometry: problems associated with the use of narrow mass precursor windows. *J Proteome Res* **14**, 700–710 (2015).
305. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J. Proteome Res.* **13**, 1234–1247 (2014).
306. Martínez-Acedo, P. *et al.* A novel strategy for global analysis of the dynamic thiol redox proteome. *Mol. Cell. Proteomics* **11**, 800–813 (2012).
307. Trevisan-Herraz, M. *et al.* SanXoT: a modular and versatile package for the quantitative analysis of high-throughput proteomics experiments. *Bioinformatics* **35**, 1594–1596 (2019).
308. Garcia-Marques, F. *et al.* A Novel Systems-Biology Algorithm for the Analysis of Coordinated Protein Responses Using Quantitative Proteomics. *Mol Cell Proteomics* **15**, 1740–1760 (2016).
309. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

310. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* **45**, D331–D338 (2017).

311. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

312. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

313. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).

314. Pérez-Rubio, P. *et al.* Extracellular vesicle depletion and UGCG overexpression mitigate the cell density effect in HEK293 cell culture transfection. *Mol. Ther. - Methods Clin. Dev.* **32**, 101190 (2024).

315. Lionaki, E. *et al.* Mitochondrial protein import determines lifespan through metabolic reprogramming and de novo serine biosynthesis. *Nat. Commun.* **13**, 1–16 (2022).

316. Lavado-García, J., Jorge, I., Cervera, L., Vázquez, J. & Gòdia, F. Multiplexed Quantitative Proteomic Analysis of HEK293 Provides Insights into Molecular Changes Associated with the Cell Density Effect, Transient Transfection, and Virus-Like Particle Production. *J. Proteome Res.* **19**, 1085–1099 (2020).

317. Colombo, M., Raposo, G. & Théry, C. Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annu. Rev. Cell Dev. Biol.* **30**, 255–289 (2014).

318. Milián, E. *et al.* Accelerated mass production of influenza virus seed stocks in HEK-293 suspension cell cultures by reverse genetics. *Vaccine.* **35**, 3423–3430 (2017).

319. Sen, D., Balakrishnan, B. & Jayandharan, G. R. Cellular unfolded protein response against viruses used in gene therapy. *Front. Microbiol.* **5**, 1–16 (2014).

320. Aponte-Ubillus, J. J. *et al.* Proteome profiling and vector yield optimization in a recombinant adeno-associated virus-producing yeast model. *Microbiologyopen* **9**, 1–12 (2020).

321. Chen, B., Retzlaff, M., Roos, T. & Frydman, J. Cellular strategies of protein quality control. *Cold Spring Harb. Perspect. Biol.* **3**, 1–14 (2011).

322. 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).

323. Hoad, M., Roby, J. A. & Forwood, J. K. Structural characterization of the porcine adeno-associated virus Po1 capsid protein binding to the nuclear trafficking protein importin alpha.

FEBS Lett. **595**, 2793–2804 (2021).

324. Bandara, L. R. & La Thangue, N. B. Adenovirus E1a prevents the retinoblastoma gene product from complexing with a cellular transcription factor. *Nature* **351**, 494–497 (1991).
325. Nemajerova, A., Talos, F., Moll, U. M. & Petrenko, O. Rb function is required for E1A-induced S-phase checkpoint activation. *Cell Death Differ.* **15**, 1440–1449 (2008).
326. Ben-Israel, H. & Kleinberger, T. Adenovirus and cell cycle control. *Front. Biosci.* **7**, 1369–1395 (2002).
327. E. Wincour, M. F. Callaham, E. H. Perturbation of the cell cycle by adeno-associated virus. *Virology* **399**, 393–399 (1988).
328. Barnes, C. R. *et al.* Genome-wide activation screens to increase adeno-associated virus production. *Mol. Ther. - Nucleic Acids* **26**, 94–103 (2021).
329. Radko, S., Jung, R., Olanubi, O. & Pelka, P. Effects of adenovirus type 5 E1A isoforms on viral replication in arrested human cells. *PLoS One* **10**, 1–18 (2015).
330. Scarrott, J. M. *et al.* Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives. *Biotechnol. J.* **18**, 1–12 (2023).
331. Brunner, S. *et al.* Cell cycle dependence of gene transfer by lipoplex, polyplex and recombinant adenovirus. *Gene Ther.* **7**, 401–407 (2000).
332. Haraguchi, T. *et al.* Transfected plasmid DNA is incorporated into the nucleus via nuclear envelope reformation at telophase. *Commun. Biol.* **5**, 1–3 (2022).
333. Verine Bä, S., Rommelaere, J., Rg, J. & Nü, P. F. Vesicular Transport of Progeny Parvovirus Particles through ER and Golgi Regulates Maturation and Cytolysis. *PLoS Pathog.* **9**, 1–16 (2013).
334. Van Niel, G. *et al.* The Tetraspanin CD63 Regulates ESCRT-Independent and -Dependent Endosomal Sorting during Melanogenesis. *Dev. Cell* **21**, 708–721 (2011).
335. Trajkovic, K. *et al.* Ceramide Triggers Budding of Exosome Vesicles into Multivesicular Endosomes. *Science.* **319**, 1244–1247 (2008).
336. Wei, D. *et al.* RAB31 marks and controls an ESCRT-independent exosome pathway. *Cell Res. 2020 312* **31**, 157–177 (2020).
337. Alenquer, M., Amorim, M. J., Graham, S. & Modis, Y. Exosome biogenesis, regulation, and function in viral infection. *Viruses* **7**, 5066–5083 (2015).
338. Ou, J. *et al.* Recent advances in upstream process development for production of recombinant adeno-associated virus. *Biotechnol. Bioeng.* **121**, 53–70 (2024).
339. Coronel, J. *et al.* Efficient Production of rAAV in a Perfusion Bioreactor Using an ELEVECTA® Stable Producer Cell Line. *Genet. Eng. Biotechnol. News* **41**, S23 (2021).

340. 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).

341. Potter, M. *et al.* A simplified purification protocol for recombinant adeno-associated virus vectors. *Mol. Ther. - Methods Clin. Dev.* **1**, 1–8 (2014).

342. 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* **28**, 147–157 (2015).

343. 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**, 16015 (2016).

344. Lavado-García, J. *et al.* Metabolic engineering of HEK293 cells to improve transient transfection and cell budding of HIV-1 virus-like particles. *Biotechnol. Bioeng.* **118**, 1649–1663 (2021).

345. Maguire, C. A. *et al.* Microvesicle-associated AAV vector as a novel gene delivery system. *Mol. Ther.* **20**, 960–971 (2012).

346. György, B., Fitzpatrick, Z., Crommentuijn, M. H. W. & Mu, D. Naturally enveloped AAC vectors for sheilding neutralising antibodies and robust gene delivery in vivo. *Biomaterials* **35**, 7598–7609 (2015).

347. Sancho-Albero, M., Medel-Martínez, A. & Martín-Duque, P. Use of exosomes as vectors to carry advanced therapies. *RSC Adv.* **10**, 23975–23987 (2020).

348. Wassmer, S. J., Carvalho, L. S., György, B., Vandenberghe, L. H. & Maguire, C. A. Exosome-associated AAV2 vector mediates robust gene delivery into the murine retina upon intravitreal injection. *Sci. Rep.* **7**, 1–10 (2017).

349. György, B. *et al.* Rescue of Hearing by Gene Delivery to Inner-Ear Hair Cells Using Exosome-Associated AAV. *Mol. Ther.* **25**, 379–391 (2017).

350. Hermida-Matsumoto, L. & Resh, M. D. Localization of human immunodeficiency virus type 1 Gag and Env at the plasma membrane by confocal imaging. *J. Virol.* **74**, 8670–9 (2000).

351. Gutiérrez-Granados, S., Cervera, L., Gòdia, F., Carrillo, J. & Segura, M. M. Development and validation of a quantitation assay for fluorescently tagged HIV-1 virus-like particles. *J. Virol. Methods* **193**, 85–95 (2013).

352. François, A. *et al.* Accurate Titration of Infectious AAV Particles Requires Measurement of Biologically Active Vector Genomes and Suitable Controls. *Mol. Ther. - Methods Clin. Dev.* **10**, 223–236 (2018).

353. Kutner, R. H., Zhang, X. Y. & Reiser, J. Production, concentration and titration of

pseudotyped HIV-1-based lentiviral vectors. *Nat. Protoc.* **4**, 495–505 (2009).

354. György, B., Fitzpatrick, Z., Crommentuijn, M. H. W., Mu, D. & Maguire, C. A. Naturally enveloped AAV vectors for shielding neutralizing antibodies and robust gene delivery in vivo. *Biomaterials* **35**, 7598–7609 (2014).
355. Lavado-García, J. *et al.* Molecular Characterization of the Coproduced Extracellular Vesicles in HEK293 during Virus-Like Particle Production. *J. Proteome Res.* **19**, 4516–4532 (2020).
356. Takahashi, A. *et al.* Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. *Nat. Commun.* **8**, 1–14 (2017).
357. Kolmus, K. *et al.* Concurrent depletion of Vps37 proteins evokes ESCRT-I destabilization and profound cellular stress responses. *J. Cell Sci.* **134**, (2021).
358. Hurwitz, S. N. *et al.* CD63 Regulates Epstein-Barr Virus LMP1 Exosomal Packaging, Enhancement of Vesicle Production, and Noncanonical NF-κB Signaling. *J. Virol.* **91**, (2017).
359. Xu, M. *et al.* The biogenesis and secretion of exosomes and multivesicular bodies (MVBs): Intercellular shuttles and implications in human diseases. *Genes Dis.* **10**, 1894–1907 (2023).
360. Russell, A. E. *et al.* Biological membranes in EV biogenesis, stability, uptake, and cargo transfer: an ISEV position paper arising from the ISEV membranes and EVs workshop. *J. Extracell. Vesicles* **8**, (2019).
361. Lorenzo, E., Miranda, L., Gòdia, F. & Cervera, L. Downstream process design for Gag HIV-1 based virus-like particles. *Biotechnol. Bioeng.* **120**, 2672–2684 (2023).
362. Graham, F. L. Growth of 293 Cells in Suspension Culture. *J. gen. Virol* **68**, 937–940 (1987).
363. Zheng, X. *et al.* Adenoviral E1a expression levels affect virus-selective replication in human cancer cells. *Cancer Biol. Ther.* **4**, 1255–1262 (2005).
364. Garcia-Moure, M. *et al.* The oncolytic adenovirus VCN-01 promotes anti-tumor effect in primitive neuroectodermal tumor models. *Sci. Rep.* **9**, 1–10 (2019).
365. Zhou, Y. *et al.* Enhanced genome editing in mammalian cells with a modified dual-fluorescent surrogate system. *Cell. Mol. Life Sci.* **73**, 2543–2563 (2016).
366. Fromigué, O., Hay, E., Barbara, A. & Marie, P. J. Essential Role of Nuclear Factor of Activated T Cells (NFAT)-mediated Wnt Signaling in Osteoblast Differentiation Induced by Strontium Ranelate. *J. Biol. Chem.* **285**, 25251–25258 (2010).
367. González-Domínguez, I., Puente-Massaguer, E., Cervera, L. & Gòdia, F. Quality assessment of virus-like particles at single particle level: A comparative study. *Viruses* **12**,

1–24 (2020).

368. Eissa, N., Kermarrec, L., Hussein, H., Bernstein, C. N. & Ghia, J. E. Appropriateness of reference genes for normalizing messenger RNA in mouse 2,4-dinitrobenzene sulfonic acid (DNBS)-induced colitis using quantitative real time PCR. *Sci. Rep.* **7**, 1–13 (2017).
369. Trivedi, P. D. *et al.* Comparison of highly pure rAAV9 vector stocks produced in suspension by PEI transfection or HSV infection reveals striking quantitative and qualitative differences. *Mol. Ther. - Methods Clin. Dev.* **24**, 154–170 (2022).
370. U.S. Department of Health and Human Services Food and Drug Administration Research. Potency testing of cell and gene therapy products. Accessed September 2023 Available at: http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/.
371. Ayuso, E. *et al.* High AAV vector purity results in serotype- and tissue-independent enhancement of transduction efficiency. *Gene Ther.* **17**, 503–510 (2010).
372. Negrini, M., Wang, G., Heuer, A., Björklund, T. & Davidsson, M. AAV Production Everywhere: A Simple, Fast, and Reliable Protocol for In-house AAV Vector Production Based on Chloroform Extraction. *Curr. Protoc. Neurosci.* **93**, 1–10 (2020).
373. Tan, E. *et al.* HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. *Front. Bioeng. Biotechnol.* **9**, 1–9 (2021).
374. Lavoie, J. N. *et al.* E4orf4, a Novel Adenovirus Death Factor That Induces p53-independent Apoptosis by a Pathway That Is Not Inhibited by zVAD-fmk volves a mechanism that includes activation of caspases. *J. Cell Biol.* **140**, 6207–6215 (1998).
375. Frisch, S. M. & Mymryk, J. S. Adenovirus-5 E1A: Paradox and paradigm. *Nat. Rev. Mol. Cell Biol.* **3**, 441–452 (2002).
376. Whyte, P., Williamson, N. M. & Harlow, E. Cellular targets for transformation by the adenovirus E1A proteins. *Cell* **56**, 67–75 (1989).
377. Fattaey, A. R., Harlow, E. & Helin, K. Independent regions of adenovirus E1A are required for binding to and dissociation of E2F-protein complexes. *Mol. Cell. Biol.* **13**, 7267–7277 (1993).
378. Ait-Si-Ali, S. *et al.* Histone acetyltransferase activity of CBP is controlled by cycle-dependent kinases and oncoprotein E1A. *Nature* **396**, 184–186 (1998).
379. Gu, B. & Wang, J. Adeno-associated virus (AAV) production. US 2022/0259572 A1_Lonza (2021).
380. Graves, D., Akkerman, N., Bachus, S. & Pelka, P. Differential Splicing of Human Adenovirus 5 E1A RNA Expressed in cis versus in trans. *J. Virol.* **95**, 1–18 (2021).

381. Hacker, D. L. & Balasubramanian, S. Recombinant protein production from stable mammalian cell lines and pools. *Curr. Opin. Struct. Biol.* **38**, 129–136 (2016).

382. Kim, T. K. & Eberwine, J. H. Mammalian cell transfection: the present and the future. *Anal. Biochem.* **397**, 3173–3178 (2010).

383. Lai, T., Yang, Y. & Ng, S. K. Advances in mammalian cell line development technologies for recombinant protein production. *Pharmaceuticals* **6**, 579–603 (2013).

384. Bandaranayake, A. D. & Almo, S. C. Recent advances in mammalian protein production. *FEBS Lett.* **588**, 253–260 (2014).

385. Howe, J. A. *et al.* Matching complementing functions of transformed cells with stable expression of selected viral genes for production of E1-deleted adenovirus vectors. *Virology* **345**, 220–230 (2006).

386. Jain, N. K., Ogden, P. J. & Church, G. M. Comprehensive mutagenesis maps the effect of all single codon mutations in the AAV2 rep gene on AAV production. *Elife* **12**, 0–32 (2023).

387. Lamartina, S. *et al.* Stringent control of gene expression in vivo by using novel doxycycline-dependent trans-activators. *Hum. Gene Ther.* **13**, 199–210 (2002).

388. Salucci, V. *et al.* Tight control of gene expression by a helper-dependent adenovirus vector carrying the rtTA2s-M2 tetracycline transactivator and repressor system. *Gene Ther.* **9**, 1415–1421 (2002).

389. Zheng, L., Baumann, U. & Reymond, J.-L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, 1–5 (2004).

390. Hoerer, M., Dubielzig, R., Ries, S., Kosfeld-Bergauer, M. Helper constructs for producing hybrid rAAV particles of various AAV serotypes. WO Patent Application 2003/074686A1. (2003).

391. Koponen, J. K. *et al.* Doxycycline-regulated lentiviral vector system with a novel reverse transactivator rtTA2S-M2 shows a tight control of gene expression in vitro and in vivo. *Gene Ther.* **10**, 459–466 (2003).

392. Urlinger, S. *et al.* Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7963–7968 (2000).

393. Zufferey, R., Donello, J. E., Trono, D. & Hope, T. J. Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors. *J. Virol.* **73**, 2886–2892 (1999).

394. Iguchi, T., Yagi, H., Wang, C.-C. & Sato, M. A Tightly Controlled Conditional Knockdown System Using the Tol2 Transposon-Mediated Technique. *PLoS One* **7**, 1–9 (2012).

395. Wang, Q. *et al.* A Robust System for Production of Superabundant VP1 Recombinant AAV Vectors. *Mol. Ther. - Methods Clin. Dev.* **7**, 146–156 (2017).

396. Muniz, L., Nicolas, E. & Trouche, D. RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. *EMBO J.* **40**, 1–21 (2021).

397. Hicks, M. J., Yang, C. R., Kotlajich, M. V. & Hertel, K. J. Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol.* **4**, 0943–0951 (2006).

398. Fong, N. *et al.* Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev.* **28**, 2663–2676 (2014).

399. David Hoggan, M., Blacklow, N. R. & Rowe, W. P. Studies of small DNA viruses found in various adenovirus preparations: Physical, biological, and immunological characteristics. *Microbiology* **55**, 1467–1474 (1966).

400. Giard, D. J. *et al.* In vitro cultivation of human tumors: Establishment of cell lines derived from a series of solid tumors. *J. Natl. Cancer Inst.* **51**, 1417–1423 (1973).

401. Longley, R. *et al.* Development of a serum-free suspension process for the production of a conditionally replicating adenovirus using A549 cells. *Cytotechnology* **49**, 161–171 (2005).

402. Siegfried, J. M. *et al.* Production of gastrin-releasing peptide by a non-small cell lung carcinoma cell line adapted to serum-free and growth factor-free conditions. *J. Biol. Chem.* **269**, 8596–8603 (1994).

403. Zhong Liu, Jr. Robert L. Longley, Marc Peter Santoro, M. V. Methods for producing A549 cell lines stable in serum-free medium suspension culture. WO2005063970A1. (2004).

404. Lester, K., Hall, M., Urquhart, K., Gahlawat, S. & Collet, B. Development of an in vitro system to measure the sensitivity to the antiviral Mx protein of fish viruses. *J. Virol. Methods* **182**, 1–8 (2012).

405. Gutiérrez-Granados, S., Cervera, L., Segura, M. de las M., Wölfel, J. & Gòdia, F. Optimized production of HIV-1 virus-like particles by transient transfection in CAP-T cells. *Appl. Microbiol. Biotechnol.* **100**, 3935–3947 (2016).

406. Fuenmayor, J., Cervera, L., Gutiérrez-Granados, S. & Gòdia, F. Transient gene expression optimization and expression vector comparison to improve HIV-1 VLP production in HEK293 cell lines. *Appl. Microbiol. Biotechnol.* **102**, 165–174 (2018).

407. Sha, J., Ghosh, M. K., Zhang, K. & Harter, M. L. E1A Interacts with Two Opposing Transcriptional Pathways To Induce Quiescent Cells into S Phase. *J. Virol.* **84**, 4050–4059 (2010).

408. Jing, X. J., Kalman-Maltese, V., Cao, X., Yang, Q. & Trempe, J. P. Inhibition of adenovirus cytotoxicity, replication, and E2a gene expression by adeno-associated virus. *Virology* **291**,

140–151 (2001).

409. T. Das, A., Tenenbaum, L. & Berkhout, B. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr. Gene Ther.* **16**, 156–167 (2016).
410. Dobbelstein, M. Replicating adenoviruses in cancer therapy. *Curr. Top. Microbiol. Immunol.* **273**, 291–334 (2004).
411. Marcos-Vadillo, E. & García-Sánchez, A. Protocol for Lipid-Mediated Transient Transfection in A549 Epithelial Lung Cell Line. *Methods Mol. Biol.* **1434**, 185–197 (2016).
412. Hubberstey AV, Pavliv M, P. R. Cancer therapy utilizing an adenoviral vector expressing only E1A. *Cancer Gene Ther.* **9**, 321–329 (2002).
413. Nagy, A. *et al.* Engineered CHO cells as a novel AAV production platform for gene therapy delivery. *Sci. Rep.* **13**, 1–16 (2023).
414. Cao, T. M., Chen, D., Barnard, G. C. & Shen, A. Recombinant adeno-associated virus production evaluation in Chinese hamster ovary cells. *Biotechnol. Bioeng.* **121**, 395–402 (2024).
415. Gao, G. P., Engdahl, R. K. & Wilson, J. M. A Cell Line for High-Yield Production of E1-Deleted Adenovirus Vectors without the Emergence of Replication-Competent Virus. *Hum. Gene Ther.* **11**, 213–219 (2004).
416. Gilbert, R. *et al.* Establishment and validation of new complementing cells for production of E1-deleted adenovirus vectors in serum-free suspension culture. *J. Virol. Methods* **208**, 177–188 (2014).
417. Liu, B. *et al.* AAV-Containing Exosomes as a Novel Vector for Improved Gene Delivery to Lung Cancer Cells. *Front. Cell Dev. Biol.* **9**, 1–12 (2021).
418. Lu, M. *et al.* Enhancing the production of recombinant adeno-associated virus in synthetic cell lines through systematic characterization. *Biotechnol. Bioeng.* **121**, 341–354 (2024).
419. Juliana, C. & Wissing, S. Method for the production of AAV (EP4006141A1). European Patent Office. (2022).
420. Hauck, B. *et al.* Undetectable transcription of cap in a clinical AAV vector: Implications for preformed capsid in immune responses. *Mol. Ther.* **17**, 144–152 (2009).
421. Fuenmayor, J. Cervera, L. Rigau, C. & Gòdia, F. Enhancement of HIV-1 VLP production using gene inhibition strategies. *Appl. Microbiol. Biotechnol.* **102**, 4477–4487 (2018).
422. Nguyen, T. *et al.* Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. *Mol. Ther. - Methods Clin. Dev.* **21**, 642–655 (2021).

ACKNOWLEDGMENTS/ AGRAÏMENTS

Aquesta tesi no hauria estat possible sense el suport de moltes persones que he tingut la sort de tenir al meu costat, a les quals vull reconèixer la seva implicació, de manera directa o indirecta durant aquesta etapa.

En primer lloc, vull agrair als meus directors, Francesc Gòdia i Laura Cervera, l'oportunitat de realitzar la tesi sota la seva guia. Gràcies Quico i Laura per confiar en mi, per les vostres paraules d'ànim i per compartir els vostres coneixements i consells des del principi fins al final d'aquesta etapa. Gràcies a vosaltres ha estat una etapa de desenvolupament professional i creixement personal molt enriquidora.

En segon lloc, vull agrair especialment el suport d'en Miquel Garcia i en Xavi León del CBATEG. Gràcies per interessar-vos genuïnament per aquest projecte i guiar-me des del principi en el món dels AAVs. Miquel, gràcies per les discussions científiques, guiar-me i aconsellar-me amb bon sentit de l'humor durant aquesta tesi. Gracies, Xavi pels teus consells i per estar sempre disposat a escoltar i respondre preguntes (que no han set poques!); els teus suggeriments i aportacions al llarg del procés i revisant la tesi han estat imprescindibles. Poder comptar amb el vostre suport científic i emocional i aprendre de vosaltres ha estat un privilegi i una sort.

També vull expressar el meu agraïment a tots els membres del grup d'Enginyeria Cel·lular i Bioprocessos, tant passats com presents, que han contribuït al llarg del procés a fer fàcils i agradables tots aquests anys. Un agraïment especial a Cristina A, Paula P, Jesús, Andy, Marc, Pol i Eli. Hem passat moments boníssims dins i fora del laboratori i ha estat un plaer compartir aquests anys amb vosaltres. Muchas gracias, Jesús, por tener la paciencia de enseñarme desde el inicio de mi máster las técnicas de biología molecular; eres un referente para mí de trabajo y generosidad; en especial también gracias por contribuir directamente a la tesis con el estudio de proteómica. Gràcies, Pol, per haver estat sempre disposat a ajudar-me, discutir resultats, donar consells i en definitiva per ser un gran amic i suport durant aquests anys. Eli, mi cubana preferida, empezamos juntas durante la pandemia y no me imagino mejor compañía para esta etapa, qué

suerte haber coincidido. També vull agrair a la Natalya i la Gaby la seva participació en la tesi com a estudiants de pràctiques de manera proactiva i entusiasta.

L'experiència no hagués estat la mateixa sense els companys del DEQBA. Vull expressar un agraïment especial als companys del grup de Lipases per la seva constant disposició a ajudar. En particular, a l'Arnau Gaset per acollir-nos al despatx, i a l'Albert i la Mariela per la seva companyia, alegria i per fer un ambient de treball molt agradable. També voldria agrair a la comissió de doctorands els bons moments organitzant les activitats. En especial, a l'Èric, gràcies per ser un amic i suport continu des de la carrera fins ara.

Per la part tècnica, vull agrair a tots els meus companys i col·laboradors de la universitat, que amb la seva ajuda i coneixements han contribuït significativament a aquesta tesi. A l'Ernest Milian, per la seva ajuda inicial amb les cèl·lules A549 i per compartir reactius i discussions amb facilitat. A la Manuela Costa i l'Arais Sierra, del servei de Citometria de l'IBB, pel seu bon humor i energia. Un agraïment especial a en Salvador Bartolomé, del servei de Bioluminiscència de la UAB, per la seva bona predisposició, suport tècnic i eficiència. També a en Martí, la Meritxell i l'Helena, del servei de Microscòpia de la UAB, i a en José Amable, del servei de NTA de l'ICMAB. També vull donar les gràcies a la Pili, la Rosi i en Manuel del Departament d'Enginyeria Química, Biològica i Ambiental pel vostre suport. A la Luisa i la Tere, per la seva alegria matinal, i a la Laura, la Rosa, la Flora i la Montse, de secretaria i administració, per estar sempre disposades a ajudar-me.

Per acabar, vull agrair el suport incondicional de la meva família i amics. A les meves amigues de tota la vida, als més recents i als que viuen a fora, per donar-me suport, pel seu positivisme i per ajudar-me a desconnectar i recuperar energies. En especial, gràcies Eloi per estar al meu costat, donar-me ànims, força i motivació.

Finalment, vull agrair el suport incondicional del meu pare, la meva mare i el meu germà. Gràcies per la vostra paciència, paraules d'ànim, consells i comprensió i per haver despertat en mi la curiositat científica des de petita. Sense vosaltres, aquesta tesi no hauria estat possible.