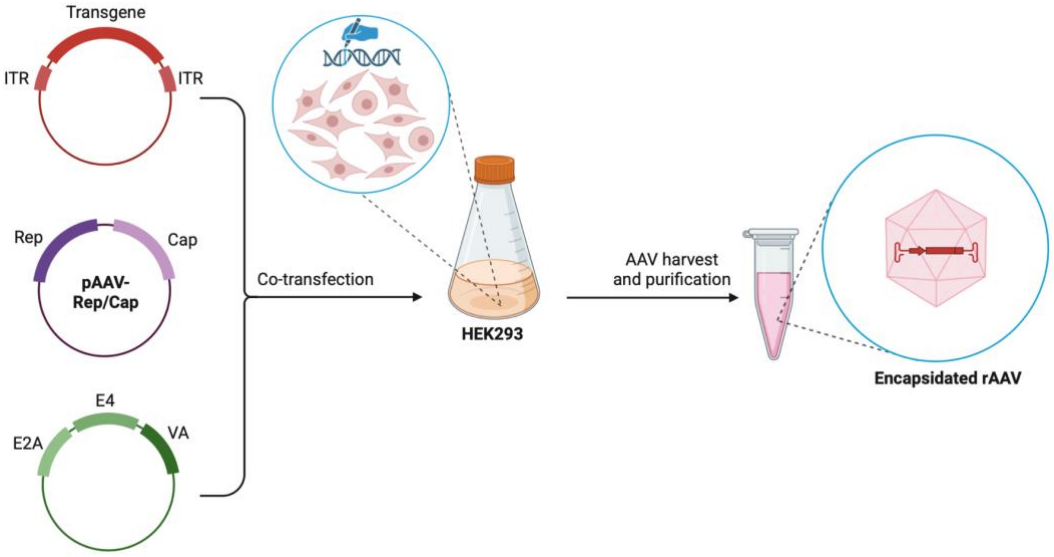
Universitat Autònoma de Barcelona

Escola d'Enginyeria

Departament d'Enginyeria Química Biològica i Ambiental

Cell Engineering Strategies in HEK293 cells to enhance rAAV production and secretion



Carlos Ruiz Ayala

Treball Final de Màster

Advisor:

Dr Laura Cervera Gracia

ABSTRACT

Recombinant adeno-associated virus (rAAV) vectors have emerged as a leading platform for in vivo gene therapies due to their favorable safety profile, broad tissue tropism, and long-lasting transgene expression. However, the high production cost of rAAV remains a major barrier to clinical scalability and patient accessibility. Transient transfection (TT) in HEK293 cells is currently the most widely used method for rAAV manufacturing, but it presents limitations in terms of scalability, batch-to-batch variability, and overall cost.

In this context, the Extended Gene Expression (EGE) protocol has demonstrated improved production yields by prolonging the expression window. Notably, an increase in rAAV titers was observed in the supernatant fraction, suggesting a greater extracellular release of viral particles. This may reduce the need for cell lysis during rAAV production, simplifying downstream processing and enhancing process efficiency.

This work explores a metabolic engineering approach to further enhance rAAV secretion by targeting host cell factors involved in extracellular vesicle (EV) biogenesis. In previous proteomic studies comparing TT and EGE conditions, a upregulation of three genes, *SNAP47*, *SMPD3*, and *VPS37B*, was observed in the EGE condition and they were selected for cell engineering due to their involvement in membrane trafficking, vesicle formation, and EV release pathways.

A HEK293 stable cell pool was generated via random integration of SNAP47 gene using antibiotic selection. Transient transfection and EGE protocols were then performed to evaluate rAAV9 production. The results showed that overexpression of *SNAP47* enhanced rAAV9 yield in the supernatant fraction, suggesting improved extracellular release. These findings support the role of vesicle-related pathways in rAAV secretion and demonstrate that targeted metabolic engineering can increase production efficiency.

Overall, this study lays the groundwork for developing stable HEK293 cell lines with enhanced rAAV secretion capacity. This strategy offers a promising route to more cost-effective and scalable rAAV manufacturing platforms, which could ultimately improve the availability of gene therapies for broader patient populations.

Key words: Gene therapy manufacturing, rAAV production, HEK293 cell engineering, metabolic engineering, stable cell line development.

ABBREVIATIONS

AAV: Adeno-Associated Virus

CNV: Copy Number Variation

CMV: Cytomegalovirus

Cq: Quantification cycle

cGMP: current Good Manufacturing Practice

ddPCR: Digital Droplet PCR

dpt: Days Post-Transfection

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

FDA: Food and Drug Administration

GFP: Green Fluorescent Protein

GOI: Gene of Interest

G418: Geneticin disulphate

HEK293: Human Embryonic Kidney 293

hpt: Hours Post-Transfection

IRES: Internal Ribosome Entry Site

ME: Medium Exchange

MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

MVB: Multi Vesicular Bodies

PBS: Phosphate-Buffered Saline

PCR: Polymerase Chain Reaction

PEI: Polyethylenimine

rAAV: Recombinant Adeno-Associated Virus

RT: Re-Transfection

SGE: Stable Gene Expression

SN: Supernatant

TGE: Transient Gene Expression

TT: Triple Transfection

Tm: Temperature of Melting

VCD: Viable Cell Density

wt: wild-type

Content

Introduction **4**

1. The gene therapy **4**

1.1. Historical context and status *4*

1.2. Viral Vectors *5*

2. Adeno-associated virus (AAV) **6**

2.1. Genome & Structure *7*

2.1.1. Adeno-associated (AAV) genome 7

2.1.2. Adenovirus helper functions 8

2.2. Adeno-associated virus life cycle *8*

2.3. Adeno-associdated virus as gene delivery vector *9*

3. Strategies for rAAV production **10**

3.1. Transient Gene Expresion (TGE) *11*

3.1.1. Batch cultures 11

3.1.2. Extended Gene Expression Protocol (EGE) 12

4. Metabolic Engineering to increase the rAAV production and egrees rate **12**

Objectives **14**

Materials and Methods **15**

Cell line and culture conditions *15*

Toxicity Assays *15*

Plasmids design for generation of the stable cell line, amplification and purification *15*

Pool generation by Random Integration *15*

Digital droplet (ddPCR) for Copy Number Variation (CNV) determination *16*

rAAV9 production by Transient Transfection and EGE *17*

Flow Cytometry *17*

Quantitative PCR (qPCR) for rAAV9 quantification *17*

Results **18**

Toxicity Assay *18*

Pool Generation *18*

ddPCR *20*

rAAV9 Production in SNAP47 Pool and WT Cells *21*

Discussion **23**

Conclusions **25**

Bibliography **26**

Introduction

1. Gene therapy

Gene therapy is defined as the administration of an active substance that contains or consists of recombinant nucleic acid, with the purpose of achieving a therapeutic, prophylactic, or diagnostic effect in humans. This effect must be directly related to the recombinant nucleic acid sequence itself or to the product resulting from its genetic expression. Unlike conventional therapies, gene therapy targets the genetic root of diseases, offering the potential for long-lasting or even curative outcomes. It represents a transformative approach in modern medicine, enabling the correction, replacement, or regulation of defective genes and the introduction of new genetic material to prevent or treat a wide range of conditions, from monogenic disorders to complex diseases[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].

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 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, directly administrated in the patient, or ex-vivo, where the target cells are extracted of the patient's body and treated with the gene therapy to finally, reintroduce them into the patient[3].

1.1. Historical context and status

In 1972, Theodore Friedmann and Richard Roblin first introduced the concept of gene therapy[3]. 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[2,3].

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[4]. Also in 2017, Luxturna® treating a congenital form of retinal dystrophy (LCA), was the first *in-vivo* AAV-based gene therapy product approved by FDA[5]. Currently, Janssen is conducting a Phase III clinical trial developing an AAV2/8-based gene therapy to treat Retinitis Pigmentosa, a disease that leads to progressive vision loss and, ultimately, complete blindness[6].

It is important to highlight that gene therapies, while potential cures for unmet medical needs, raise significant concerns regarding patient access. For instance, Hemgenix®, an AAV gene therapy encoding Factor IX for Hemophilia B bleeding disorder, is priced at $3.5M per dose. 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[7,8].

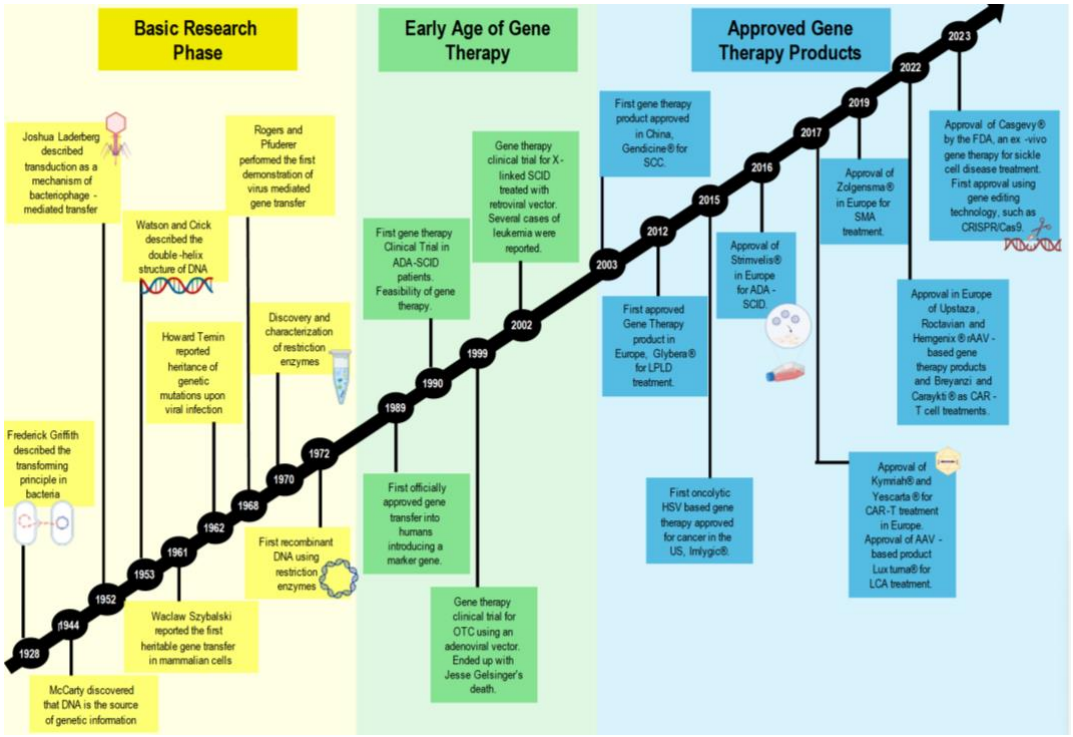


Figure 1. Timeline and key milestones in gene therapy development. Figure from[9]

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. However, the scope of these products in terms of the potencial patient population that can be reached, is limited due to their remarkably high prices, constantly breaking price records. For example, Hemgenix is the most recently approved product with potencial of replacing the Hemophilia B lifelong therapies to a single infusion. Although at a price of 3.5$ million, making it the world's most expensive therapy.

These elevated costs are partly due to recoup the R&D and production expenses. 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.

For example, Roctavian® for Hemophilia A, administered through intravenous infusion requires 6.0×10^{13} vg/kg, so approximately 4.5×10^{15} vg per patient need to be injected. Production scale is estimated based on a productivity of 5×10^{14} vg/L and a downstream recovery of 20%. So, for one dose is necessary 450L, skyrocketing the total production scale[10,11].

The exceptionally high cost of these therapies poses substantial barriers to accessibility. 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[12].

1.2. 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. Different vectors are used for gene therapy with different characteristics. These can be classified into viral or non-viral vectors[13].

Non-viral vectors include nucleic acids covered by lipid nanoparticles (LNPs), exosomes, cationic polymers, inorganic nanoparticles and polymer hydrogels to enhance its stability. 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. LNPs are nowadays more popular after its use in several COVID vaccines[14].

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%. Viral vectors are engineered to lack wild-type 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[15,16].

They are further categorized based on their capacity for integration into the host genome into integrating and non-integrating vectors. 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[13].

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[17]. Over time, AAVs have become the preferred choice for viral gene therapy applications due to their favourable characteristics compared to other viral vectors. Despite their limited capacity for carrying genetic material compared to adenoviral vectors (4.4 kb)[17].

2. Adeno-associated virus (AAV)

AAVs were discovered by Atchinson *et al*. in 1965 as a contaminant virus in wild-type adenovirus preparations[17]. Their replication was found to be dependent on the presence of adenoviruses, hence named adeno-associated viruses[18,19].

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. This facilitated the development of AAVs as a new type of genetically modified viral particle, capable of carrying therapeutic DNA[20,21].

Today, recombinant AAVs (rAAVs) are considered the preferred choice in the gene therapy field over other vectors for their desirable characteristics. These include their non-pathogenic nature, 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[22].

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. To date, no evidence directly links rAAV use to tumorigenesis in clinical settings[23,24].

2.1. Genome & Structure

AAvs are formed by a single-stranded DNA (ssDNA genome) in an icosahedral capsid. The AAV genome is organized in three open reading frames, as can be observed in **Figure 2**, which encode for structural and regulatory viral proteins. AAV genome of 4.7kb is flanked at both ends by 145 nucleotide long T-shaped hairpin-like structures called inverted terminal repeats (ITR)[25]. The ITRs play a crucial role in AAV life cycle, providing self-priming activity for DNA replication and packaging signal to encapsulate the ssDNA into the capsid[20,26].

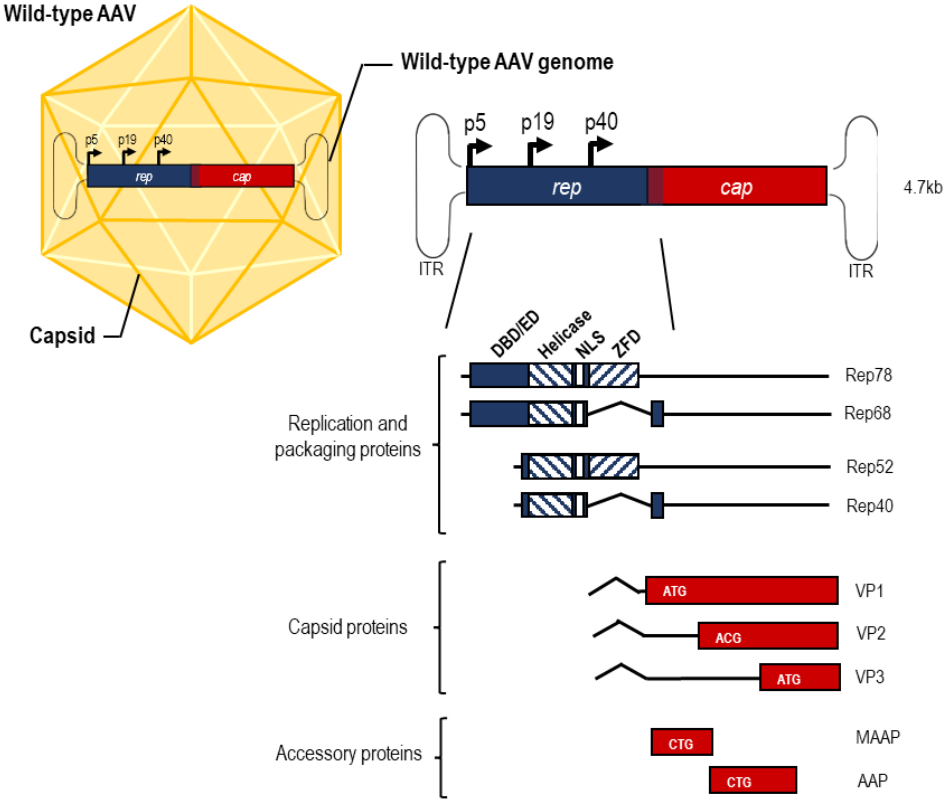


Figure 2. Wild-type AAV genome structure. Figure from[9]

2.1.1. Adeno-associated (AAV) genome

The large Rep proteins (Rep78 and Rep68) are transcribed from the p5 promoter. Both Rep78 and Rep68 proteins exhibit DNA binding, helicase and endonuclease activities. Their sequence-specific DNA binding domain drives AAV DNA replication by binding to the ITR at the Rep binding site (RBS)[27–29]. Similarly, small Rep proteins (Rep52 and Rep40) are expressed from the downstream p19 promoter encoded within Rep ORF. They are essential for the efficient packaging of the AAV genome into capsid structures via the ATP-dependent helicase activity and for ssDNA accumulation during replication[30].

Transcription of the *cap* gene is controlled by the p40 promoter encoded, encoding the capsid proteins VP1, VP2 and VP3. 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[31].

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)[32]. 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[33].

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[34].

Additional non-structural accessory proteins, expanding the coding capacity of the AAV genome have been identified through alternate ORF within the *cap* gene. The presence of AAP has been shown to increase capsid protein stability.

Also, 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. Therefore, the association of MAAP with EVs has been shown to alter AAV egress and kinetics, promoting the extracellular secretion of particles[35].

2.1.2. Adenovirus helper functions

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[36,37].

The set of adenoviral genes essential for AAV2 replication and propagation consists of *E1A, E1B, VA RNAs, E2A* and *E4orf6*. First, *E1A* upregulates transcription from the AAV p5 promoter, which controls the expression of large Rep protein. Additionally, E1A transactivates early adenovirus promoters, including *E1B, the VA RNAs, E2A, and E4orf6*[38–40].

In the context of adenovirus-infected cells, the *E1B* gene products (19K and 55K proteins) exhibit anti-apoptotic functions presumably to counteract E1A pro-apoptotic properties and ensure productive infection and cellular transformation[41–43].

DBP plays a pivotal role in increasing intracellular levels of AAV genome, stimulates AAV DNA replication and capsid protein production[31]. Finally, virus-associated RNA (VA RNAs) are noncoding RNAs involved in enhancing viral protein translation by circumventing host cell's antiviral response[44].

2.2. Adeno-associated virus life cycle

The AAV life cycle begins with cell recognition and attachment of the capsid to negatively charged glycans, followed by endocytosis via plasma membrane invagination. For AAV2, this occurs efficiently within 30-60 min[45]. After endosomal escape, AAV particles travel to the nuclear periphery and enters through the nuclear pore complex (NPC) as intact particles, mediated by the canonical nuclear import pathway, involving the VP1/VP2 N-terminal nuclear localization signal (NLS) and importin-β. Once inside the nucleus, viral particles are disassembled, and the viral genome is released in a process not fully known[46].

In the presence of adenovirus helper functions, AAV DNA replication proceeds. As shown in the **Figure 3B**, replication process relies on the self-annealing property of the ITRs, which provides a free 3'OH end acting as an origin for viral replication. Capsid protein oligomerize in the cytoplasm and capsid assembly occurs in the nucleolus. The newly synthesized ssDNA is packaged into capsids by the helicase/ATPase domain of the small Rep proteins[47].

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. Additionally, several studies have identified that a fraction of rAAV isolated from the extracellular medium are associated with exosomes, termed exo-rAAV[48,49]. These exo-rAAVs show enhanced transduction efficiency capacity[50]. Notably, the MAAP protein, reported as a viral egress factor, promoted rAAV exit from the cells by interacting with the surface of extracellular vesicles. 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[35]. A schematic representation of the AVV replication is shown in **Figure 3A**.

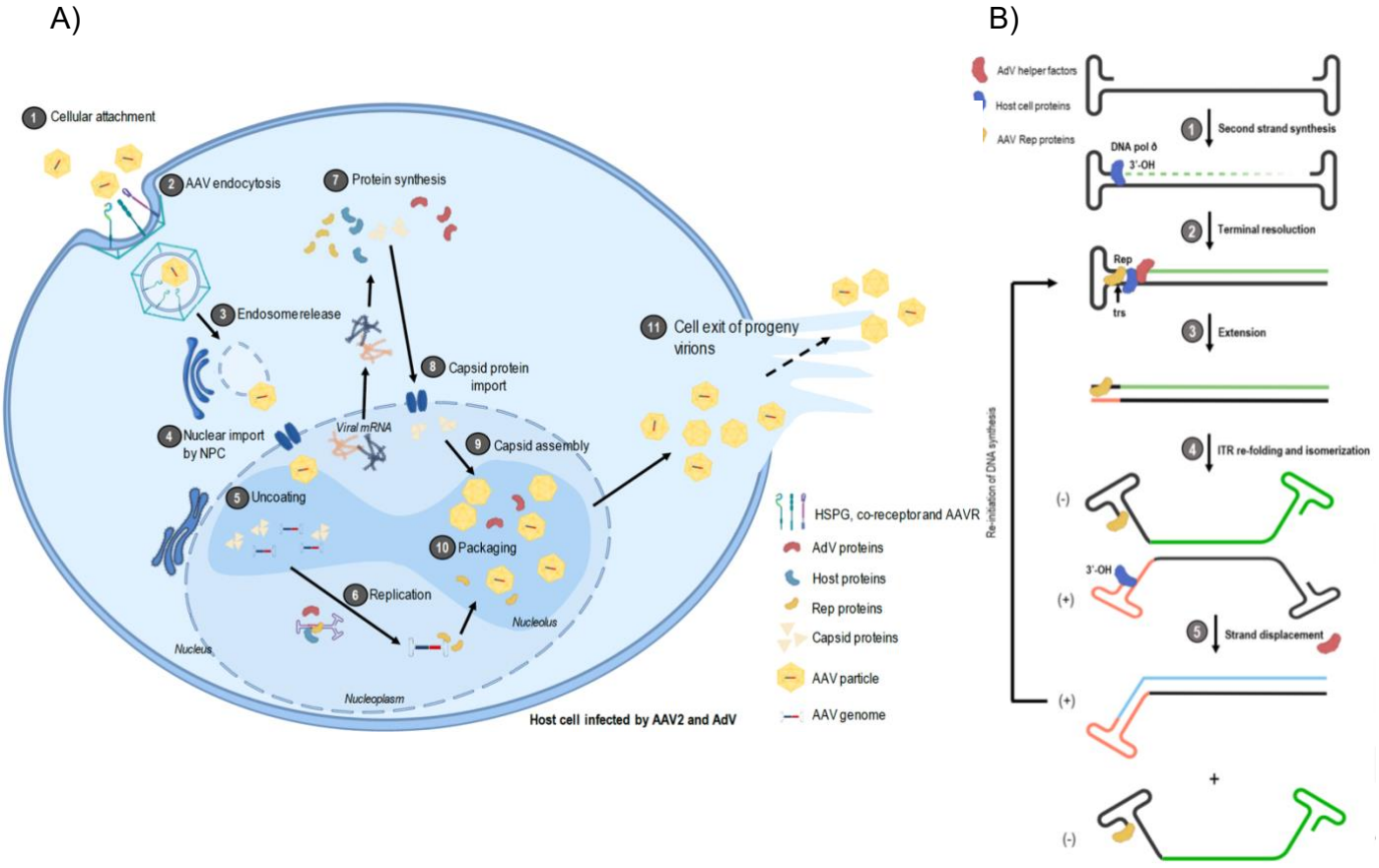


Figure 3. A) Schematic representation of AAV2 replication by AdV co-infection. B) AAV DNA replication. Figure adapted from[9]

2.3. Adeno-associated virus as gene delivery vector

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 4A**). 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 4B,** pGOI_ITR). The optimal packaging capacity is between 4.1-4.9 kb[51] .

The CMV and CAG promoters are commonly employed in clinical trials across different therapeutic areas for their robust constitutive expression. In contrast, tissue-specific promoters are preferred for blood diseases to minimize off-target activity[22].

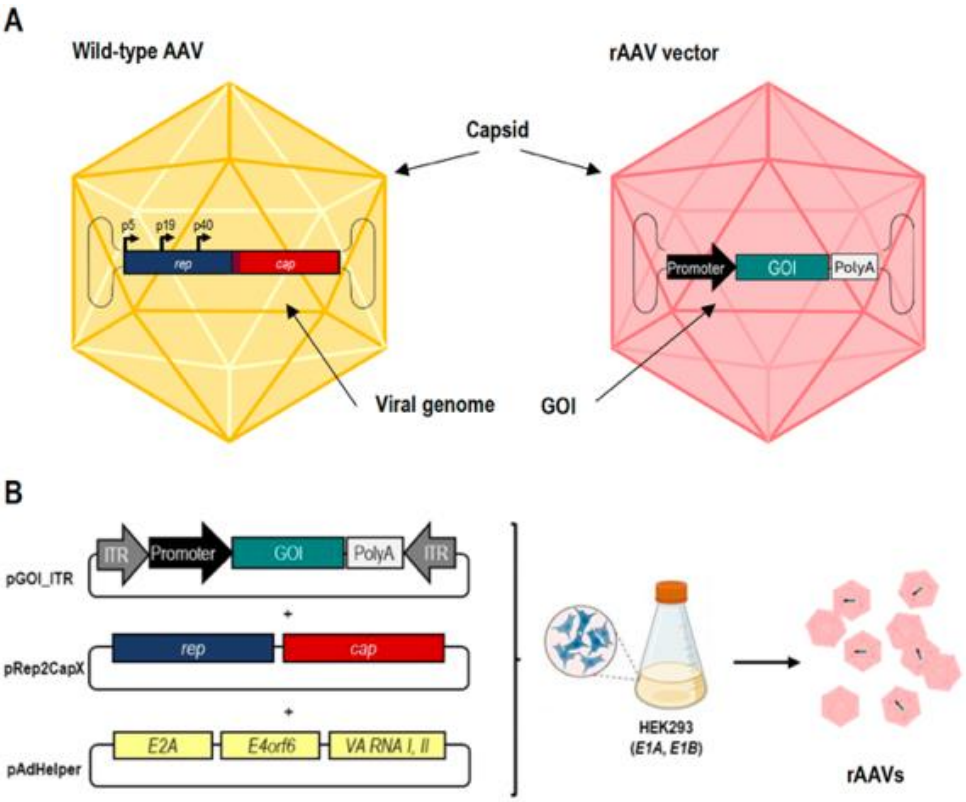


Figure 4. **AAV vectorization**. A) The viral wt AAV genes (rep and cap) flanked by ITRs are replaced with the therapeutic gene of interest (GOI) in recombinant AAV (rAAV) genome. B) Plasmids used for rAAV production into HEK293 cells by transient transfection method. Figure from[9]

As can be observed in **Figure 4B,** for rAAV production, the most common method is triple transfection into HEK293 cells. One plasmid contains the GOI gene, another the AVV genes (rep gene and cap gene), and finally, the last one contains AdV helper gene set (E2A, E4 and VA RNAs). These cells were originally transformed with the E1 gene, resulting in constitutive expression of E1A and E1B proteins, essential for AVV replication[52]. It is common practice to use the *cap* gene from one serotype while utilizing the Rep proteins and ITRs from AAV2 (pRep2CapX).

3. Strategies for rAAV production

Two different approaches are identified based on the gene expression method: transient gene expression (TGE) and stable gene expression (SGE). 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 involve 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[10].

In TGE, DNA is not integrated into the host cell's genome, thereby it is lost over time due to cell division and environmental factors, limiting production period to 72 hpt[53]. In addition, there are several transfection-associated disadvantages. First, delivery of multiple plasmids, 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 scalability[54]. Overall, the cost, 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[54]. This method avoids the use of high amounts of plasmid 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. On the other hand, 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[55]. Due to this, the manufacturing of AAVs using SGE has not yet been established at the required levels for clinical applications.

3.1. Transient Gene Expresion (TGE)

TGE in HEK293 cells through transient triple transfection (TT) using PEI as the transfection agent is the most widely used strategy for rAAV production. This method has been disclosed to be the production method for commercial rAAV-based products such as Luxturna® [5].

As represented in the **Figure 4B** three plasmids are co-transfected: pGOI_ITR plasmid containing 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. 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[48,54]. Althought some bioreactors, such as iCellis®, enable industrial-scale production using adherent cells[56].

Consequently, the transitioning of HEK293 cells adapted to grow in suspension and in serum-free culture media marked a significant advancement for rAAV production, facilitating a more homogeneous and controllable rAAV manufacturing processes.

3.1.1. Batch cultures

Smith *et al.* in 2003 reported rAAV production by TT into HEK293 cells in suspension conditions using Erlenmeyer flasks. For scaling up production, large shake flasks offer a straightforward option, with reported working volumes ranging from 200 mL to 1 L[53]. 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[57]. 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[58]. In addition, a decrease in cell-specific productivity is observed when cells are transfected at high cell density cultures (HCD), known as the cell density effect (CDE)[53]. 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. Therefore, the scale-up of batch operation mode is limited by cell density at the time of transfection, limiting volumetric productivity[59].

3.1.2. Extended Gene Expression Protocol (EGE)

The Extended Gene Expression (EGE) protocol, developed by Cervera et al., has emerged as a promising tool for enhancing viral particle production. This protocol, developed to produce virus like particles (VLPs) 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[60]. 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.

In AAV production, the EGE protocol has demonstrated a notable advantage by enabling the collection of AAVs directly from the culture supernatant, thereby eliminating the need for cell lysis. This is particularly significant, as cell lysis is a major source of batch-to-batch variability in viral vector manufacturing.

By avoiding lysis, the purification process is also simplified, due to the lower initial levels of host cell DNA and host cell proteins. This, in turn, reduces the reliance on expensive endonucleases, ultimately helping to lower the overall production costs. Furthermore, implementation of the EGE protocol for rAAV production has resulted in a remarkable threefold increase in total rAAV yield from the supernatant, compared to traditional triple transfection (TT) methods using batch production and lysate recovery[61].

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. 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[62]. However, the EGE method requires a high amount of GMP-grade plasmids for repeated transfection rounds, which substantially increases the costs of the process.

4. Metabolic Engineering to increase the rAAV production and egrees rate

A major challenge in rAAV manufacturing is the high cost and complexity of downstream processing. Currently, rAAV vectors are predominantly harvested through cell lysis, as approximately 50–70% of the vectors remain intracellular, making cell disruption essential. This process not only requires extensive purification steps but also leads to significant yield losses [63]. Enhancing the natural release of rAAVs into the culture supernatant would enable direct recovery from the extracellular fraction, simplifying purification and reducing overall production costs. Therefore, a key objective of this work is to increase the proportion of rAAVs secreted into the supernatant, potentially eliminating the need for cell lysis. In this context, host cell engineering strategies that promote extracellular viral egress represent a promising approach.

In previous studies, rAAV production and secretion was enhanced using the EGE protocol, as shown in **Figure 6A**. A proteomic analysis was performed to investigate the dynamics of the host cell proteome during rAAV production in HEK293 cells and to compare the metabolic differences between the TT and EGE production methods. The results revealed that vesicle transport pathways were upregulated under the EGE condition compared to the TT 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.

As can be observed in the **Figure 5**, the formation starts 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[64].

At least three routes have been described: the ESCRT pathway, the tetraspanin pathway and a lipid-based route involving lipids such as sphingomyelin. First, the increased expression of VPS37B in the EGE condition suggests the involvement of an ESCRT-I mediated sorting mechanism[65]. 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. Ceramides, inducing negative membrane curvature, lead to the formation of ILVs as the neutral sphingomyelinase SMPD3[9,65]. 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)[64], heavily involved in this process, and SNAP47, showing the highest upregulation in EGE compared to TT.

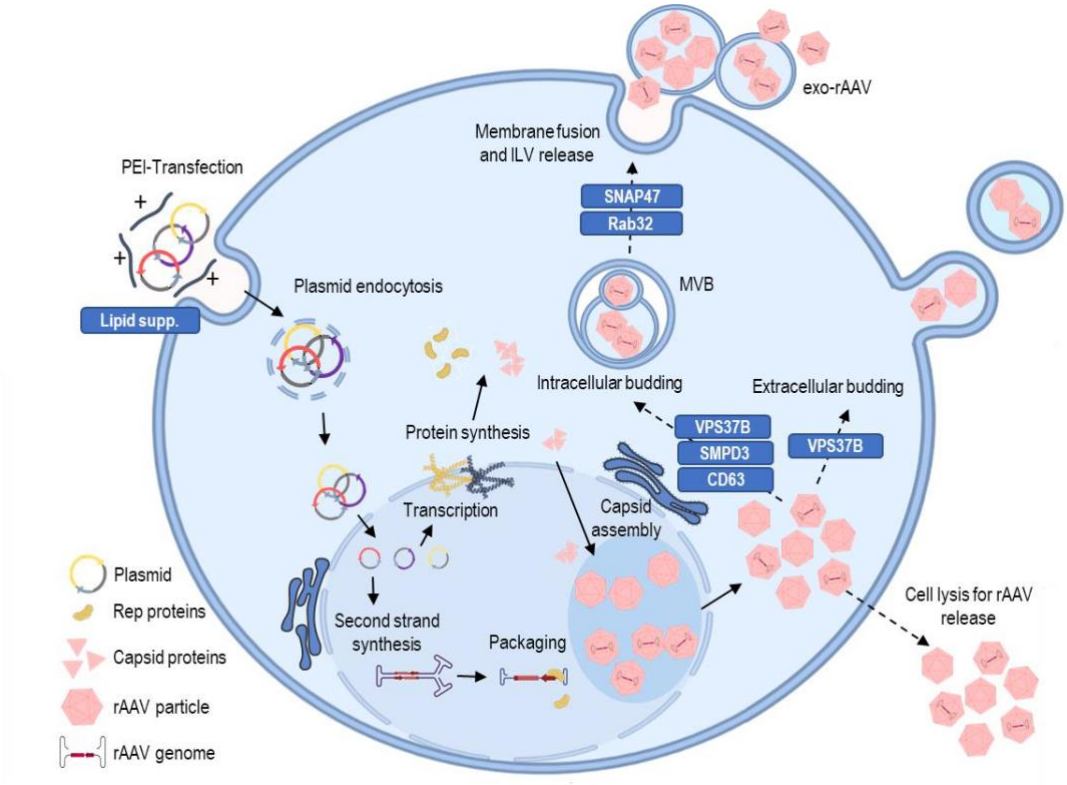


Figure 5. **Schematic representation of the rAAV production pathway by triple transfection.** Figure from[9]

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. Next, these previously identified gene candidates affecting the EV production were co-transfected with TT plasmids to analyze its effect on exo-rAAV production. The results (**Figure 6B)** show that the overexpression of SNAP47, SMPD3 and VPS37B enhances the total rAAV9 total yield and more importantly, the percentage of rAAV9 found in supernatant fraction.

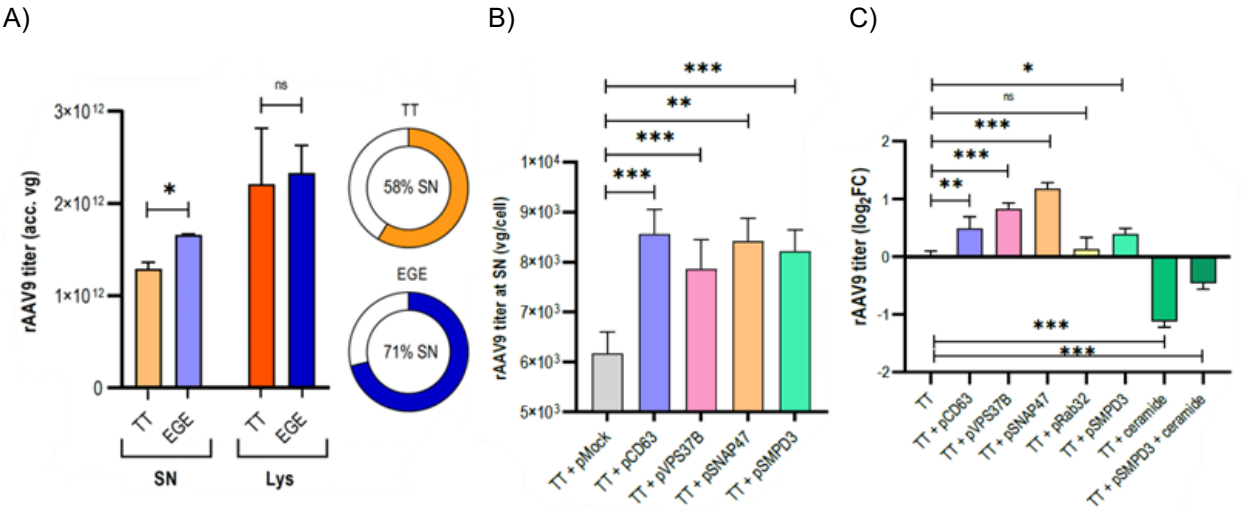


Figure 6. Characterization of rAAV9 titer at extracellular media and produced exo-rAAVs. A) rAAV9 titer of quadruple transfection using standard triple transfection plasmids adding one of the candidate genes. B) rAAV9 titer of triple transfection plasmids co-transfected with one candidate gene. C) rAAV9 titer expressed as log2FC (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. Data was analyzed by ANOVA test for significant difference *p<0.05; **p<0.01; ***p<0.001. Figure adapted from[9].

Metabolic engineering represents a promising strategy to overcome current limitations in rAAV production and secretion. In this context, the generation of a stable HEK293 cell line with high expression of the *SNAP47*, *SMPD3*, and *VPS37B* genes could be beneficial for improving extracellular rAAV production, increasing yield and recovery in the cell supernatant fraction using the TT production method.

This innovative approach will not only allow functional evaluation of these molecular modulators in optimizing viral vector production but also lay the groundwork for developing more efficient, scalable, and cost-effective cell platforms. Collectively, this work aims to contribute to reducing manufacturing costs and improving the accessibility of rAAV-based gene therapies, facilitating their industrial production and clinical application.

Objectives

- Generate a HEK293 cell line which contains extra-copies of *SNAP47*, *SMPD3*, and *VPS37B* genes integrated by random mutation.

- Study if these HEK293 cell line enhance the rAVV harvest in the supernatant fraction and the rAVV production.

Materials and Methods

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 150 rpm in a LT-X Kuhner shaker (LT-X Kuhner, Birsfelden, Switzerland).

Cell culture was FreeStyle TM F17 Expression media (Gibco, Thermofisher Scientific, San Jose, CA, USA) supplemented with 8mM Glutamax^{TM} (Gibco, TermoFisher Scientific, San Jose, CA, USA), 0.1% Pluronic^{TM} 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 50ug/L. Cell concentration and viability were determined using the Nucleocounter® NC-3000 automatic cell counter (Chemometec, Allerod, Denmark) according to manufacurer's instructions.

Toxicity Assays

100 µl of HEK293 cells were seeded in a 96-well plate at $0.5 \cdot 10^6$ cells/ml together with 10 µl of increasing antibiotic concentrations, all points in triplicate. The plate was cultured under standard HEK293 culture conditions. After two days, 20 µl of MTT reagent (Promega, WI, USA) was added to each well and the plate was incubated for one hour at 37 ºC. Just before the MTT addition, a standard curve of cell concentrations from 0 to $3 \cdot 10^6$ viable cells/ml was seeded Absorbance at 490 nm was read on Multilabel Plater Reader VICTOR3® (Perkin Elmer, MA, USA).

Plasmids design for generation of the stable cell line, amplification and purification

Plasmids were designed based on a promoter trap strategy, which consists of placing the gene of interest (GOI) followed by an internal ribosome entry site (IRES) and a resistance gene. This configuration allows for co-expression of the GOI and the selectable marker from a single transcript, ensuring that antibiotic-resistant clones have also integrated and are expressing the GOI[66,67]. Three plasmids were constructed: one encoding SNAP47 with G418 as the selectable marker, another encoding SMPD3 with Zeocin, and a third encoding VPS37B with Hygromycin.

Plasmids were amplified in Escherichia coli Top10 strain grown in LB medium at 36ºC supplemented with 100 µg/ml ampicillin or Kanamycin (Sigma, St. Louis, MO, USA). Plasmid purification was carried out using the ZymoPure II Plasmid Maxiprep Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer"s instructions.

Pool generation by Random Integration

Transfections were carried out at a cell density of 2×10^6 cells/mL or 1×10^6 cells/mL in Erlenmeyer flasks, using a total culture volume of 20 mL and a final DNA concentration of 1µg/mL or 0,5 µg/mL. 25 kDa linear polyethylenimine (PEI Max, PolySciences, Warrington, PA, USA) was used as transfection reagent. DNA/PEI complexes were formed by adding PEI to plasmid DNA diluted in medium (10% of the total culture volume).

PEI was added at a 1:2 ratio (DNA/PEI), once done, the mixture is incubated for 15 min at room temperature and then introduced to the cell culture. All conditions were performed in triplicate[53]. Transfection was performed using circular or linearized DNA. Linearization of the plasmids were achieved by FastDigest BshTI (ThermoFisher, Waltham, MA, USA) restriction, according to manufacturer recommendations.

Antibiotic selection with Geneticin, Zeocin, or Hygromycin B is initiated 48 hours post-transfection to favour random plasmid integration into the host genome. This process often relies on the non-homologous end joining (NHEJ) DNA repair pathway, which requires time to incorporate foreign DNA into chromosomal breaks. This delay allows sufficient time for the integration and expression of the resistance gene.

Digital droplet (ddPCR) for Copy Number Variation (CNV) determination

Genomic DNA was purified using Wizard Genomic DNA Purification Kit (Promega, WI, Madison, USA) according to manufacturer"s instructions. Purity and concentration of the extracted genomic DNA was assessed with a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Two pairs of primers were designed to target and amplify fragments of 122 and 141 bp from SNAP47 and Ribonuclease P Protein Subunit P30 (RPP30) genes respectively, using the online PrimerQuest Tool from IDT (Integrated DNA Technologies IDT-DNA, Leuven, Belgium).

Table1: Primers to target and amplify RPP30 and SNAP47 genes

Gene	Forward Primer	Reverse Primer
RPP30	5"-CCAACCTCATGCCACCCAGACCATC-3"	5"-CCCGCGCTAGGAATCAGACC AACAC-3"
SNAP47	5"-ATGAGCAGGGATGTCTGC-3"	5"-GTGCTGTCAGTCATG AACCT-3"

To optimize the annealing temperatures of the primers to be used in the ddPCR, a qPCR with a temperature gradient of 55.4 to 60.7 ºC was performed a CFX384 Touch Real-Time PCR Detection System. The cycling conditions were as follows: 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.

Each 10 µl PCR mixture contained 2x iTaq Universal SYBR Green (#1725151, Bio-Rad, Hercules, CA, USA), primers at a final concentration of 500 nM (IDT-DNA, Leuven, Belgium) and 100 ng of gDNA. Based on this assessment, the annealing temperature was determined to be 58 ºC for both RPP30 and SNAP47 primers.

Each ddPCR mixture contained 2x EvaGreen ddPCR Supermix (#1864034, Bio-Rad, Hercules, CA, USA), the corresponding pair of primers at a final concentration of 500 nM (IDT-DNA, Leuven, Belgium) and 10 ng of gDNA. The 20 µl PCR mixtures were loaded into DG8 Cartridges together with 70 µl of Droplet Generation Oil (#1863009 and #1863005 respectively, Bio-Rad, Hercules, CA, USA). After gasket attachment, the cartridges were placed into the QX200 Droplet Generator (#1864002 Bio-Rad, Hercules, CA, USA), which generated ~20.000 droplets per sample. The content was then transferred to a 96-well PCR plate, sealed by means of a PX1 PCR Plate Sealer (#1814000, Bio-Rad, Hercules, CA, USA) and cycled using the following conditions: 95 °C, 5 min + 40x[95 °C, 30 s + 58°C, 60 s] + 4 °C, 5 min + 90 °C, 5 min + 4 °C hold. The QX200 Droplet Reader (#1864003, Bio-Rad, Hercules, CA, USA) was used to read the individual cycled droplets.

rAAV9 production by Transient Transfection and EGE

For rAAV production, transient transfection (TT) was carried out in HEK293 cells using equimolar amounts of pHelper, pGFP_ITR, and pRep2Cap9 plasmids. Cells were transfected at 2 × 10^6 cells/mL in 20 mL cultures using 1 µg/mL total DNA. DNA/PEI complexes were prepared by mixing plasmid DNA (10% culture volume) with 25 kDa linear PEI (PEI Max, PolySciences, Warrington, PA, USA) at a 1:2 DNA:PEI ratio, incubated for 15 minutes at room temperature, then added to the culture[53]. For EGE method, media exchange was performed at 48 hpt by centrifuging the cell culture at 300xg for 5 minutes. The supernatant was harvested, and the cell pellet was resuspended with fresh medium. After that, re-transfection (RT) is performed following the same TT protocol.

Flow Cytometry

The percentage of eGFP-positive cells was assessed using a Beckman Coulter CytoFLEX LX cytometer (Beckman Coulter, Brea, CA, USA). 5,000 events were analysed in each sample. First, Side Scatter Height (SSC-H) vs. Forward Scatter Area (FSC-A) and FSC-A vs. Forward Scatter Height (FSC-H) density plots were used to gate HEK 293 cells. Afterwards, the percentage of GFP-positive cells was assessed in a GFP FITC-A vs. APC-A density plot. Data was processed using CytExpert v.2.3 software (Beckman Coulter Inc., Brea, CA, USA).

Quantitative PCR (qPCR) for rAAV9 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. 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 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. Thermocycling profile was based on manufacturer's recommendations in a CFX384 Touch Real-Time PCR Detection System: 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

For the establishment of a standard curve, a gBlock (IDT, Leuven, Belgium) containing the targeted GFP region of a known concentration was used. 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.

Table2: Primers to target and amplify GFP and gBlock sequence which contains the targeted GFP region.

Gene	Forward Primer	Reverse Primer
GFP	5"-AGCTGGACGGCGACGTAA -3"	5"- GTGCAGATGAACTTCAGGGTCA-3"
gBlock	5"-AGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA AGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCA TCTGCACCACCGGCAAGCT GCCCGTGCCCTGGCCCACCCTCGTGAC-3"	

Results

Toxicity Assay

The three constructed plasmids were designed to generate a stable cell line using a promoter trap strategy. In this design, the gene of interest (GOI) is placed upstream of an internal ribosome entry site (IRES) followed by an antibiotic resistance gene. This configuration enables expression only if the construct integrates downstream of an active endogenous promoter. Specifically, the plasmids encoded *SNAP47* (with G418 resistance), *SMPD3* (with zeocin resistance), and *VPS37B* (with hygromycin B resistance), respectively.

To generate the random stable cell line, cellular pools were subjected to selection with G418, zeocin and hygromycin. As the antibiotic effect and tolerance differs between cell lines, toxicity assays were performed to assess the HEK293-3F6 cellular response to each antibiotic and further determine its minimal selective concentration for effective selection.

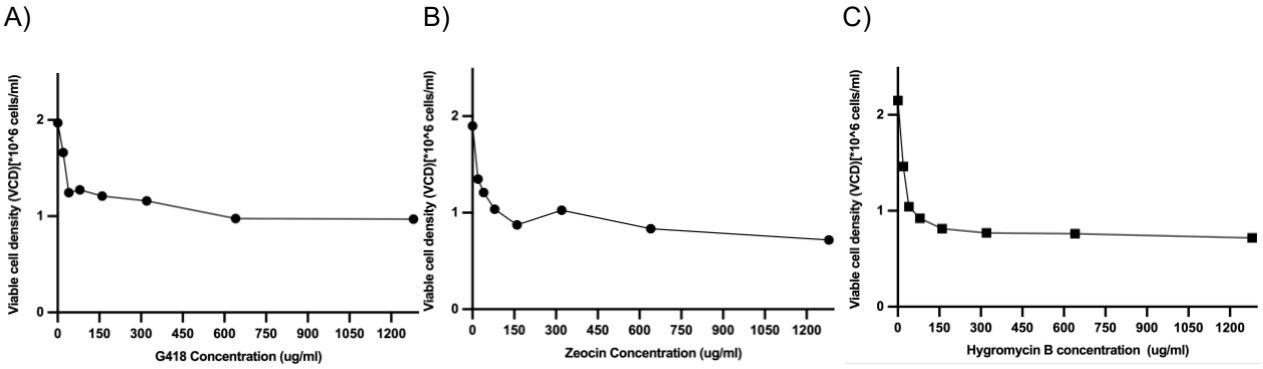


Figure 7. MTT toxicity assay results. A) Viable cell density after being cultured for 48 hours under different concentrations of geneticin (G418). B) Viable cell density after being cultured for 48 hours under different concentrations of Zeocin C) Viable cell density after being cultured for 48 hours under different concentrations of Hygromycin B.

A concentration of 40 µg/mL for Geneticin, 80 µg/mL for Zeocin, and 160 µg/mL for Hygromycin B were established as the minimal effective concentrations for selection in this specific HEK293 cell line.

Pool Generation

To generate the stable cell line producing plasmids were transfected as circular plasmid. If the DNA is not randomly integrated, it is diluted over time after each cell division. Therefore, after several days under selection, the surviving cells will be the ones that spontaneously integrated the plasmid in their genome.

Initially, no selective pressure was added to the cell culture in order to allow the recovery of the transfected culture and the expression of the resistance genes by the cell machinery. At 2 days post-transfection (dpt), cells were supplemented with the corresponding selective antibiotics: G418 for *SNAP47*, zeocin for *SMPD3*, and hygromycin B for *VPS37B*. All antibiotics were added at a final concentration of 225 µg/mL.

As can be observed in **Figure 8**, 5 dpt a significant decrease in cell viability was expected due to the death of the non-transfected population. However, no such effect was observed under G418 selection. Consequently, the concentration of G418 was increased to 450 µg/mL.

Following this adjustment, cell proliferation ceased, and a notable decline in cell viability was observed, indicating effective selective pressure.

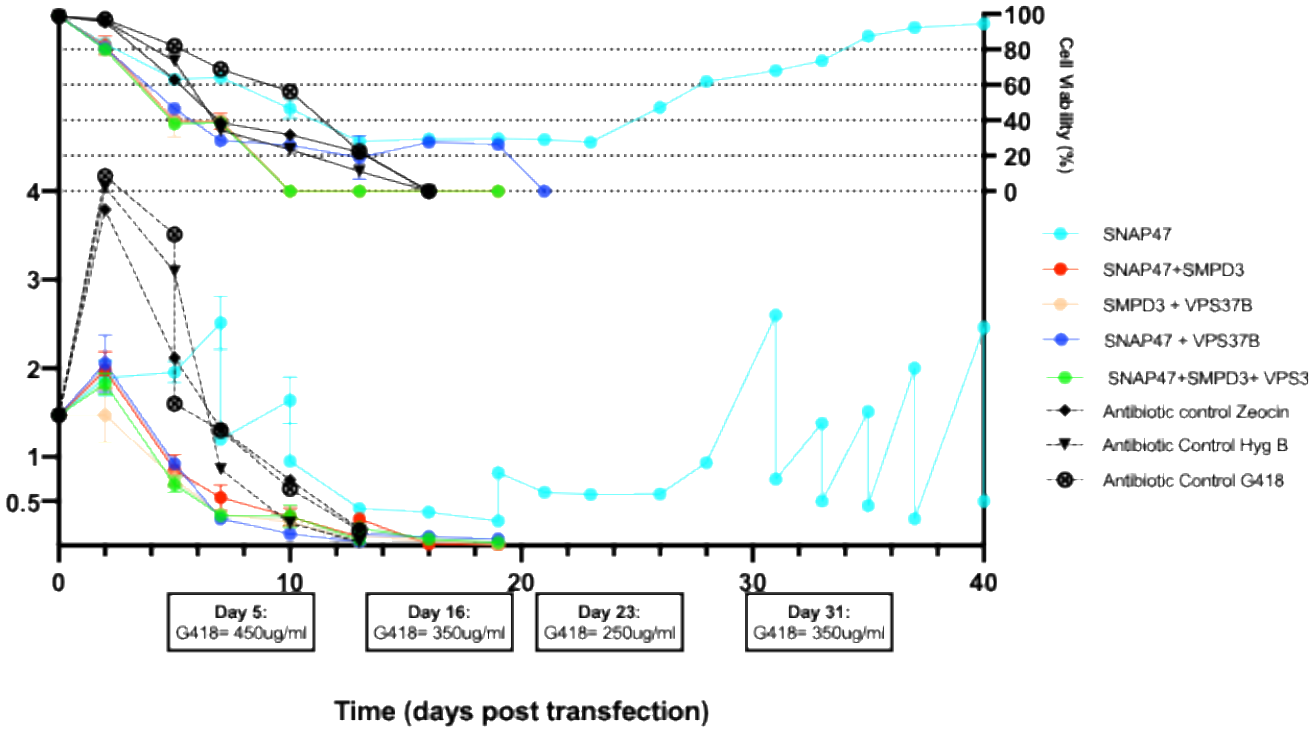


Figure 8. Cellular pool evolution. Viable cell density and cell viability. After each cell count, medium exchange and antibiotic was realized. Means are represented by dots and error bars indicates standard deviations.

The viability drop experienced by cultures from 2 dpt to 14 dpt was caused by the death of the non-transfected population, the loss of transgene expression in transfected cells due to plasmid dilution and degradation, and the inability of some transfected cells to integrate the construct into a genomic region with sufficient transcriptional activity to allow expression of the antibiotic resistance gene via the FMDV IRES sequence. The condition where cells were only transfected with *SNAP47* plasmid was able to survive the selection

At 16 dpt, the concentration of G418 was reduced, allowing the surviving population to recover. From day 22 dpt to 28 dpt, the positive selection of those cells was traduced in an increase of viability. However, pools were not showing cell growth. At 30 dpt, viability and growth was totally recovered.

After generating the stable pool containing extra copies of *SNAP47*, cells were subsequently transfected with the *SMPD3* plasmid, that was previously linearlized, and subjected to dual selection with G418 and zeocin, aiming to obtain a stable cell pool co-expressing both *SNAP47* and *SMPD3*. Additionally, the pool was exposed to G418 and zeocin simultaneously, as well as to zeocin alone, to evaluate whether the combination of both antibiotics could exert a synergistic cytotoxic effect leading to cell death.

Finally, cells were cultured in the absence of antibiotics to assess whether the removal of selective pressure would impact the viability and proliferation of the stable pool. On day 9, G418 was reintroduced to these cultures to evaluate whether the temporary removal of selection affected cell viability or growth dynamics.

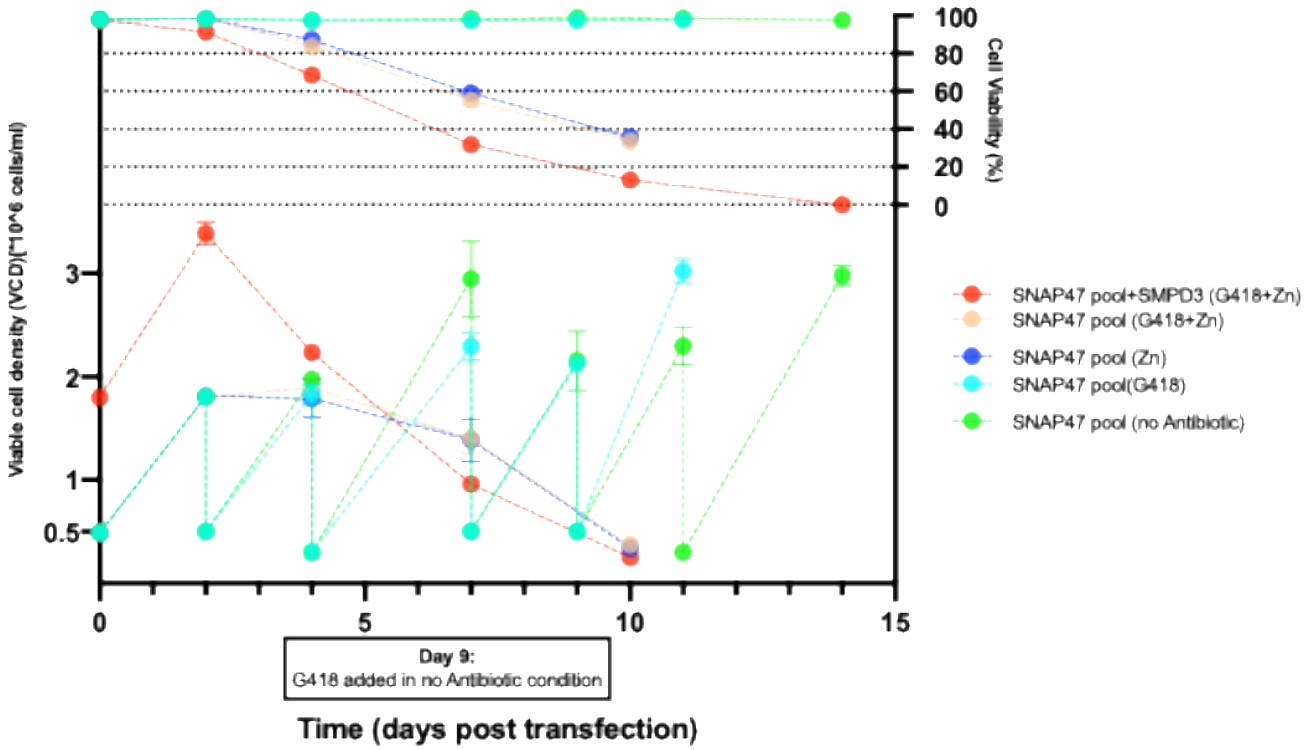


Figure 9. Cellular pool evolution. Viable cell density and cell viability. After each cell count, medium exchange and antibiotic was realized. Means are represented by dots and error bars indicates standard deviations.

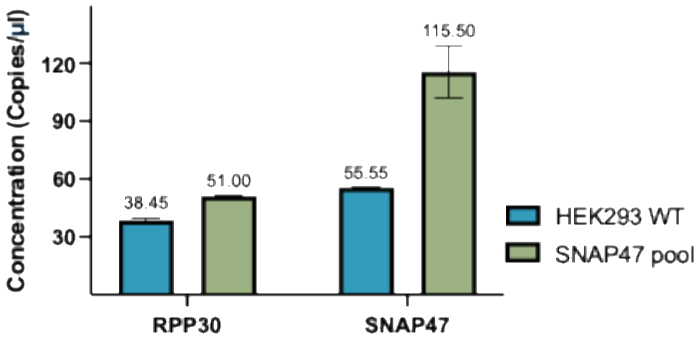
The decrease in cell viability observed between 2 and 10 days post-transfection (dpt) was attributed to several factors: the death of the non-transfected population, the loss of transgene expression in transfected cells due to plasmid dilution and degradation, and the inability of some transfected cells to integrate the construct into a genomic region with sufficient transcriptional activity to allow expression of the antibiotic resistance gene via the FMDV IRES sequence. The cell pool transfected with the *SMPD3* plasmid did not survive zeocin selection. No differences in viability were observed between the condition subjected to dual antibiotic selection (G418 and zeocin) and the condition exposed to zeocin alone. Finally, the addition of G418 to the *SNAP47* pool previously cultured without antibiotics did not affect cell growth or viability.

ddPCR

Copy number variation (CNV) is a phenomenon in which the presence of number of copies of a particular DNA segment varies relative to a reference genome. To fully characterize the *SNAP47* pool generated. This was assessed using digital droplet PCR (ddPCR), an ultrasensitive method for absolute DNA quantification overcoming inherent limitations of qPCR and microarray techniques when analysing CNV. In this approach, PCR reactions for each condition—containing the exact same amount of diluted genomic DNA, were partitioned into over 20,000 nanoliter, sized droplets. Then, they were PCR amplified until the reaction end-point and each droplet was individually analysed to determine whether it was positive (contained the target amplicon) or negative (did not).

This resulted in the determination of the copy number for each studied construct. The Ribonuclease P Protein Subunit P30 (RPP30) gene from wild-type HEK293 was used as reference assay, as it has been reported to present two copies in human diploid genomes[68].

A)



B)

	HEK293 WT	*SNAP47* pool
RPP30 (copies/µl)	38.45±1.20	51±0.42
SNAP47 (copies/ul)	55.55±0.35	115,5±13.43
Ratio *SNAP47/RPP30*	1.44±0.04	2.26±0.26
SNAP47 copies per genome	3	4

Figure 10. Absolute quantification of *RPP30* and *SNAP47* genes by ddPCR. A) Copy number of RPP30 and *SNAP47* genes in HEK293-derived cell lines. The RPP30 gene, known to exist in two copies per diploid genome, was used as a reference to estimate the number of *SNAP47* copies integrated per cell.

As shown in Figure 10, wild-type HEK293 cells exhibited three copies of *SNAP47* per genome. In contrast, the selected stable cell pool displayed an average of four copies per genome, suggesting that, on average, one additional copy of *SNAP47* was successfully integrated. It is important to note that this result corresponds to a heterogeneous cell pool, and therefore the copy number reflects the mean across a mixed population rather than a clonal cell line.

rAAV9 Production in *SNAP47* Pool and WT Cells

Cell cultures are seeded at 0.5 × 10^6 cells/mL and transfected upon reaching a density of 2 × 10^6 cells/mL. Two different rAAV9 production strategies are compared: the standard triple transfection (TT) protocol under batch conditions, and an enhanced protocol based on Extended Gene Expression (EGE), which includes a medium exchange (ME) and a re-transfection (RT) step at 48 hpt, as illustrated in **Figure 11A**. Both production strategies are applied to wild-type HEK293 cells, used as control, and to the previously generated HEK293 stable cell pool carrying additional copies of *SNAP47* (from now on SNAP47 cell pool).

As shown in **Figure 11B**, following transfection, all cultures remain at a viable cell density of approximately 2 × 10^6 cells/mL, with no further increase observed. A decline in viability is detected for all conditions; however, at 96 hpt, cells cultured using the EGE protocol exhibit approximately 10% higher viability compared to those under the TT protocol. The percentage of transfected cells is measured during all the process by the percentage of GFP-positive (GFP^+) cells. As can be observed in **Figure 11C,** a higher proportion of GFP^+ cells are observed at 72 hpt and 96 hpt in cultures subjected to the EGE protocol compared to those transfected using the TT protocol.

Finally, as can be observed in **Figure 11D,** significant difference in rAAV production is observed, under TT protocol between wild-type HEK293 cells and the stable SNAP47 cell pool. More specifically, significant differences are observed in the number of viral genomes (vg) found in the supernatant when using Transient Transfection (TT) production method. However, no significant differences are observed in the total vg or when the Extended Gene Expression (EGE) method was used for production. On the other hand, significant difference are found in the total vg (AAVs harvested in the supernatant and inside the cells, harvested by lysis) at 96 hpt between the TT and EGE production method.

Moreover, in **Figure 11E**, shows that a higher proportion of rAAV9 particles is detected in the supernatant of the stable SNAP47 cell pool compared to HEK293 WT cells applying TT protocol as production method. On the other hand, EGE allows for higher recovery of AAVs in the cell supernatant for WT and SNAP47 cell pool.

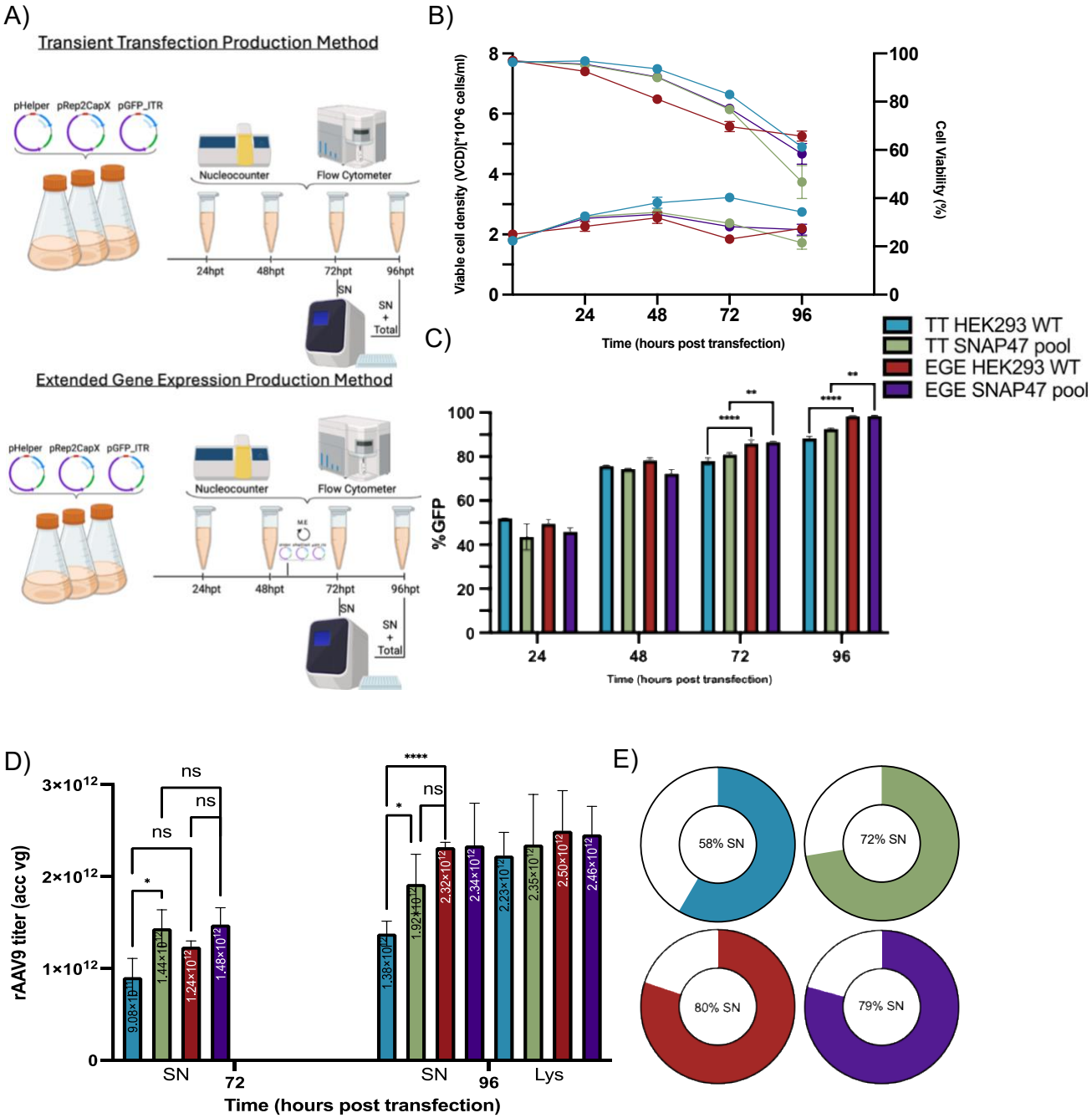


Figure 11. **Experimental design and cell culture analysis during rAAV9 production.** A) Experimental Workflow B) Viable cell density and viability analyzed during production time course C) Flow cytometry analysis of transfection efficiency based on GFP-positive cell percentage. D) rAAV9 production results at 72 hpt and 96 hpt using TT and EGE production methods E) Percentage of rAAV9 in the supernatant at 96 hpt. Data was analyzed by two-way analysis of variance (ANOVA) for significant difference *p<0.05; **p<0.01; ***p<0.001; ****p<0,0001 error bars indicate the SD.

Discussion

The generation of a stable cell line through the simultaneous transfection of three genes does not appear to be an effective strategy. As shown in **Figure 8**, no cell survival is observed under triple antibiotic selection. A similar outcome is observed under double transfection conditions, in which no viable cell population is successfully selected under dual antibiotic pressure. These results suggest that sequential integration, introducing one gene at a time, may be a more viable approach, despite being more time-consuming, as the development of a stable cell pool requires 20-30 days.

However, as shown in **Figure 9,** challenges remain even when using this sequential strategy**.** In the case of SMPD3, whose plasmid was linearized to facilitate genomic integration, selection with zeocin does not result in a surviving cell population. This suggests that successful integration and cell survival depend on additional factors beyond vector design and antibiotic selection. It also raises the possibility that the promoter trap vector strategy may not be the most effective approach for generating stable cell lines.

The promoter trap strategy is commonly used to identify endogenous promoters through the random integration of a reporter gene. Reporter Expression occurs only when the construct integrates downstream of an active endogenous promoter, thereby enabling the discovery and characterization of transcriptionally active genomic regions[66]. This strategy is particularly attractive because it can reduce the risk of transgene silencing, which often occurs when strong promoters are used to generate stable cell lines expressing proteins of interest, such as in CHO cells for antibody production[69,70].

However, further studies are required to evaluate the suitability and efficiency of promoter trap designs for the generation of stable cell lines via random genomic integration. In parallel, alternative strategies should be explored, such as the use of promoters with a lower risk of transgene silencing or more effective IRES sequences to ensure robust expression of the selectable marker.

Additionally, it is possible that alternative vector designs could enable successful double or even triple transfections under multiple antibiotic selection. Although integrating one gene at a time is more time-consuming, this strategy allows for the selection of cell clones with higher expression levels of each individual gene at each step. Therefore, despite requiring longer timelines, sequential integration may be more effective in generating a stable cell line with strong overexpression of all three genes, reducing the probability that any of them is sub-optimally integrated or expressed.

Despite multiple attempts to generate a stable cell pool expressing all three genes, the successful generation of a stable cell pool with extra copies of *SNAP47* was confirmed, as supported by ddPCR results and the cell culture progression shown in **Figure 8**. By day 28 of selection, the cell pool exhibited good viability and growth parameters under G418 selection pressure, suggesting that the insert is integrated into a transcriptionally active genomic region. Digital droplet PCR analysis further indicates that the heterogeneous cell pool contains approximately one additional copy of *SNAP47* per genome, as demonstrated in **Figure 10.**

This represents a significant achievement, as stable genomic integration of even a single gene through random integration and antibiotic selection is a complex and time-consuming process, and it provides a solid foundation for further engineering efforts.

The next steps for this pool should include characterization of *SNAP47* protein expression compared to wild-type cells, followed by the isolation of a clone exhibiting the highest expression levels of SNAP47. This could be assessed by dot-blot or western blot analysis.

Although *SNAP47* protein expression has not been yet characterized in the stable cell pool, the results are consistent with previous findings shown in **Figure 6**. The stable cell pool with extra copies of *SNAP47* exhibits similar outcomes to those observed when *SNAP47* was co-transfected during Transient Transfection (TT). As shown in **Figure 6B**, *SNAP47* increased the number of AAV particles in the supernatant by approximately 30% compared to the control condition. Similarly, **Figure 11D** shows that at 96 hours post-transfection (hpt), AAV production in the supernatant increase by approximately 30% when using the stable cell pool. Moreover, this trend is also reflected in the **Figure 11E**, where the percentage of rAAV9 recovered from the supernatant increase from 58% in wild-type cells to 72% when is produced using the stable cell pool under TT protocol.

Moreover, a significant difference in total rAAV9 production between wild-type HEK293 cells and the stable *SNAP47* pool was expected based on the results observed in **Figure 6C**. However, this difference was not observed. It is possible that *SNAP47* expression is sufficient to enhance vector egress but not to increase overall production. Additionally, it is important to note that the results presented in **Figure 6** were obtained under transient *SNAP47* expression, which may elicit different cellular responses compared to stable overexpression in the engineered cell pool.

On the other hand, as previously shown in **Figure 6A**, the EGE protocol does not increase the total amount of rAAV9 produced. Although, in **Figure 11D,** significant differences are observed in total viral genome (vg) titters between TT and EGE protocol.

Finally, an additional unexpected result is the decline in cell viability under EGE conditions observed in **Figure 11B**. It was initially anticipated that no significant difference would occur between 72 hpt and 96 hpt; however, a drop in viability was observed. Nonetheless, this decrease is less pronounced than the observed under the TT protocol. These findings suggest that further studies and optimizations are required to refine the EGE protocol for efficient rAAV9 production and to maintain high cell viability, allowing for an extended production window. Despite the drop in viability, the effect of medium exchange (ME) and re-transfection (RT) at 48 hpt was clearly reflected in the increased percentage of GFP-positive cells at 72 and 96 hpt, as well as in the higher AAV production compared to the TT protocol, as shown in **Figure 11C**.

Conclusions

This study aimed to enhance rAAV9 production and secretion by engineering a HEK293 cell line through stable integration of host genes involved in vesicle trafficking: *SNAP47, SMPD3*, and *VPS37B*. These genes were selected based on proteomic data obtained under EGE production conditions, where extracellular vesicle pathways were found to be upregulated.

A key achievement of this work was the successful generation of a stable HEK293 cell pool containing extra copies of *SNAP47*. ddPCR analysis confirmed the presence of approximately one additional copy of *SNAP47* per genome. Importantly, this pool maintained high viability under G418 selection, indicating stable integration in a transcriptionally active region. Although *SNAP47* protein expression remains to be quantified, the phenotypic impact of its overexpression was evident.

Despite multiple attempts, it was not possible to generate stable pools co-expressing more than one gene. The failure to establish pools expressing both *SNAP47* and *SMPD3*, or triple-expressing cells, suggests that additional factors such as vector design, promoter efficiency, or selection strategy may significantly influence success. Although sequential gene integration is proposed as a potential solution to overcome the limitations of dual and triple antibiotic selection, its effectiveness remains to be fully validated. Future studies should explore improved vector architectures, alternative expression systems, or site-specific integration approaches to enable stable multi-gene expression.

When compared to wild-type HEK293 cells, the *SNAP47*-enriched pool showed enhanced release of rAAV9 particles into the supernatant following transient transfection. Under the TT protocol, a significant increase of approximately 30% in supernatant viral genomes was observed, consistent with previous results using transient *SNAP47* co-transfection. However, total rAAV9 production did not differ significantly, suggesting that *SNAP47* enhances extracellular release rather than intracellular synthesis. Notably, a higher proportion of total rAAV9 was recovered in the supernatant fraction (72%) compared to wild-type cells (58%).

Regarding production protocols, EGE showed higher transfection efficiency and better cell viability at 96 hpt than TT. EGE method demonstrated clear advantages over the standard TT protocol. Significant increase in total rAAV9 titers. Notably, up to 80% of the total rAAV9 particles were recovered in the supernatant under EGE conditions, supporting its superior efficiency in promoting extracellular viral release.

In conclusion, this work demonstrates that metabolic engineering of HEK293 cells via stable *SNAP47* integration enhances rAAV9 extracellular release, offering a promising strategy to facilitate downstream processing. Although further work is required to achieve multi-gene integration and optimize expression systems, the results provide a strong foundation for the development of more efficient, scalable, and cost-effective rAAV production platforms.

Bibliography

1. *Directive 2001/83/EC 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. *Biomedicine & Pharmacotherapy* **153**, 1–20 (2022).
3. Friedmann, T. & Roblin, R. Gene Therapy for Human Genetic Disease? *Science (1979)* **175**, 949–955 (1972).
4. Ylä-Herttuala, S. Endgame: Glybera Finally Recommended for Approval as the First Gene Therapy Drug in the European Union. *Molecular Therapy* **20**, 1831–1832 (2012).
5. 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* **Volume 13**, 179–208 (2020).
6. Byrne, B. J. *et al.* Current clinical applications of AAV-mediated gene therapy. *Molecular Therapy* **33**, 2479–2516 (2025).
7. Ned Pagliarulo. Pricey new gene therapies for sickle cell pose access test. *BiopharmaDive* (2023).
8. Heo, Y.-A. Etranacogene Dezaparvovec: First Approval. *Drugs* **83**, 347–352 (2023).
9. Laia Bosch Molist. Enhancing recombinant adeno-associated virus (rAAV) production through cell and gene engineering. (Escola d'Enginyeria UAB, 2024).
10. Sha, S. *et al.* Cellular pathways of recombinant adeno-associated virus production for gene therapy. *Biotechnol Adv* **49**, 107764 (2021).
11. 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).
12. De Luca, M. & Cossu, G. Cost and availability of novel cell and gene therapies. *EMBO Rep* **24**, 2–5 (2023).
13. Wang, D. & Gao, G. State-of-the-art human gene therapy: part I. Gene delivery technologies. *Discov Med* **18**, 67–77 (2014).
14. Ramamoorth, M. Non Viral Vectors in Gene Therapy- An Overview. *Clinical and diagnostic research* **9**, 1–6 (2015).
15. 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).
16. 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).
17. Bouard, D., Alazard-Dany, N. & Cosset, F. Viral vectors: from virology to transgene expression. *Br J Pharmacol* **157**, 153–165 (2009).

18. Atchison, R. W., Casto, B. C. & Hammon, W. McD. Adenovirus-Associated Defective Virus Particles. *Science (1979)* **149**, 754–756 (1965).

19. 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).

20. 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. *Proceedings of the National Academy of Sciences* **79**, 2077–2081 (1982).

21. 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).

22. Au, H. K. E., Isalan, M. & Mielcarek, M. Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings. *Front Med (Lausanne)* **8**, 1–14 (2022).

23. Li, C. & Samulski, R. J. Engineering adeno-associated virus vectors for gene therapy. *Nat Rev Genet* **21**, 255–272 (2020).

24. Smith, R. H. Adeno-associated virus integration: virus versus vector. *Gene Ther* **15**, 817–822 (2008).

25. 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).

26. 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).

27. 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).

28. 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).

29. Hong, G., Ward, P. & Berns, K. I. In vitro replication of adeno-associated virus DNA. *Proceedings of the National Academy of Sciences* **89**, 4673–4677 (1992).

30. King, J. A. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J* **20**, 3282–3291 (2001).

31. 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).

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

33. 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).

34. Flotte, T. R. & Berns, K. I. Adeno-Associated Virus: A Ubiquitous Commensal of Mammals. *Hum Gene Ther* **16**, 401–407 (2005).

35. Elmore, Z. C. *et al.* The membrane associated accessory protein is an adeno-associated viral egress factor. *Nat Commun* **12**, 1–13 (2021).

36. Meier, A. F., Fraefel, C. & Seyffert, M. The Interplay between Adeno-Associated Virus and Its Helper Viruses. *Viruses* **12**, 8–12 (2020).

37. 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).

38. 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).

39. Matsushita, T. *et al.* The adenovirus E1A and E1B19K genes provide a helper function for transfection-based adeno-associated virus vector production. *Journal of General Virology* **85**, 2209–2214 (2004).

40. 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).

41. Rao, L. *et al.* The adenovirus E1A proteins induce apoptosis, which is inhibited by the E1B 19-kDa and Bcl-2 proteins. *Proceedings of the National Academy of Sciences* **89**, 7742–7746 (1992).

42. White, E. Regulation of the cell cycle and apoptosis by the oncogenes of adenovirus. *Oncogene* **20**, 7836–7846 (2001).

43. White, E. Function of the adenovirus E1B oncogene in infected and transformed cells. *Semin Virol* **5**, 341–348 (1994).

44. Vachon, V. K. & Conn, G. L. Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Res* **212**, 39–52 (2016).

45. Bartlett, J. S., Wilcher, R. & Samulski, R. J. Infectious Entry Pathway of Adeno-Associated Virus and Adeno-Associated Virus Vectors. *J Virol* **74**, 2777–2785 (2000).

46. 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).

47. Charman, M. & Weitzman, M. D. Replication Compartments of DNA Viruses in the Nucleus: Location, Location, Location. *Viruses* **12**, 1–20 (2020).

48. Lock, M. *et al.* Rapid, Simple, and Versatile Manufacturing of Recombinant Adeno-Associated Viral Vectors at Scale. *Hum Gene Ther* **21**, 1259–1271 (2010).

49. György, B. & Maguire, C. A. Extracellular vesicles: nature's nanoparticles for improving gene transfer with adeno-associated virus vectors. *WIREs Nanomedicine and Nanobiotechnology* **10**, 1–13 (2018).

50. Hudry, E. *et al.* Exosome-associated AAV vector as a robust and convenient neuroscience tool. *Gene Ther* **23**, 380–392 (2016).

51. Dong, J.-Y., Fan, P.-D. & Frizzell, R. A. Quantitative Analysis of the Packaging Capacity of Recombinant Adeno-Associated Virus. *Hum Gene Ther* **7**, 2101–2112 (1996).

52. 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).

53. 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).

54. 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).

55. Yang, Q., Chen, F. & Trempe, J. P. Characterization of cell lines that inducibly express the adeno-associated virus Rep proteins. *J Virol* **68**, 4847–56 (1994).

56. Lennaertz, A., Knowles, S., Drugmand, J.-C. & Castillo, J. Viral vector production in the integrity® iCELLis® single-use fixed-bed bioreactor, from bench-scale to industrial scale. *BMC Proc* **7**, 59 (2013).

57. 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).

58. van der Loo, J. C. M. & Wright, J. F. Progress and challenges in viral vector manufacturing. *Hum Mol Genet* **25**, R42–R52 (2016).

59. 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).

60. 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).

61. Godia-Casablancas, F. B.-T. M.-F. G.-M. M. C.-G. L. L.-M. X. M.-L. M. G.-G. S. Methods for the manufacture of recombinant viral vectors (EP2020058527). . (2020).

62. Fuenmayor, J., Cervera, L., Gòdia, F. & Kamen, A. Extended gene expression for Gag VLP production achieved at bioreactor scale. *Journal of Chemical Technology & Biotechnology* **94**, 302–308 (2019).

63. Jungbauer, A. & Wheelwright, S. Downstream processing of AAV based gene therapy vectors. *Sep Purif Technol* **368**, (2025).

64. 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).

65. Alenquer, M. & Amorim, M. Exosome Biogenesis, Regulation, and Function in Viral Infection. *Viruses* **7**, 5066–5083 (2015).

66. Chen, J. *et al.* Identification of a novel endogenous regulatory element in Chinese hamster ovary cells by promoter trap. *J Biotechnol* **167**, 255–261 (2013).

67. Chaudhary, S., Pak, J. E., Gruswitz, F., Sharma, V. & Stroud, R. M. Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. *Nat Protoc* **7**, 453–66 (2012).

68. Dyavar, S. R. *et al.* Normalization of cell associated antiretroviral drug concentrations with a novel RPP30 droplet digital PCR assay. *Sci Rep* **8**, 3626 (2018).

69. Škalamera, D. *et al.* Generation of a Genome Scale Lentiviral Vector Library for EF1α Promoter-Driven Expression of Human ORFs and Identification of Human Genes Affecting Viral Titer. *PLoS One* **7**, (2012).

70. Zeh, N., Schmidt, M., Schulz, P. & Fischer, S. The new frontier in CHO cell line development: From random to targeted transgene integration technologies. *Biotechnol Adv* **75**, (2024).