

**Developing recombinant Adeno-Associated Virus producer
cell lines through systems synthetic biology**

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Min Lu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Professor Wei-Shou Hu

September 2023

© Min Lu 2023

Acknowledgements

I want to express my greatest gratitude to my advisor, Professor Wei-Shou Hu. Words alone cannot adequately describe how grateful I am to have the chance of being his student and growing under his guidance and mentorship. Professor Hu always considers the training of students as his top priority. I am always inspired by his comprehensive knowledge in bioengineering, intuition regarding new trends of biotechnology, and well-established collaboration and connection across academia and industry. My fruitful PhD journey has been more than I could have ever imagined. Professor Hu has not only guided me to become a well-rounded scientist but also a better human being, ready to contribute to the society.

I would like to express my gratitude to all our collaborators for their insightful perspectives and feedback on our projects, as well as for the recourses they provided to facilitate our research. I thank Professor Daniel Schmidt and Professor George Aslanidi, who offered informative critiques based on their expertise in AAV biology, pointed out details we may have overlooked, and guided our projects towards success. I appreciate Professor Nikunj Somia and Professor Scott McIvor for their valuable insights into molecular biology, particularly regarding the construction of gene expression cassette. And I would like to express my appreciation to Professor Ben Hackel, Professor Casim Sarkar, and Professor George Aslanidi for serving on my committee.

Next, I would like to thank the colleagues and alumni of the Hu group who have played an important role in guiding me to become an outstanding scientist and have brought me unforgettable experiences during my time at the University of Minnesota. I truly appreciate all the profound positive impact that Dr. Zion Lee has had on me. He has taught

me a wide range of knowledge about molecular biology and experimental skills. He also has been my mentor, assisting me in adapting to the new environment and developing critical thinking abilities. Most importantly, he is a friend and a big brother who always shares his experiences and offers advice to help me become a better person. Special thanks to Dr. Thu Phan, who has been a role model for me, showing me how to adapt as an international student and providing valuable insights into my research, presentations, communication skills, and job applications. Tremendous thanks to Yen-An Lu and Janani Narayan, who have shared all the pleasant and challenging time together with me during my PhD journey. I also thank all the members of the Hu group members and alumni for all the support and joyful times we shared together: Dr. Chris Stach, Dr. Conor O'Brien, Dr. Sofie O'Brien, Dr. Jen One, Dr. Kevin Ortiz-Rivera, Dr. Hansol Kim, Dr. Meghan McCann, Dr. Qian Ye, Wen Cai, Dr. Yu-Chieh Lin, and Dr. Pedram Motallebnejad.

A huge thanks to the undergraduates who have dedicated their time and effort to our projects, making the success possible. I would like to express my gratitude to Ibrahim Irfanullah for contributing his intelligence to building a kinetic model and helping us achieve accomplishment. I also want to thank Daniela Batres, Joshua Sereno, and Annika Bowen for their relentless efforts on the projects. Being a mentor to you is both my honor and one of the most precious experiences in my PhD journey.

Finally, I would like to express my deepest appreciation to my family and my girlfriend. I thank my girlfriend for her emotional support, as she talks to me over the phone every day, even though we are far apart, so I can share all the good and the bad things. I am grateful to my parents and sister for their full support regarding my decision to study abroad and for their assistance in all aspects during this long journey.

Abstract

Recombinant adeno-associated virus (rAAV) is one of the most widely used vehicles for gene therapy due to its effectiveness in *in vivo* gene delivery, long-term persistency, and safety. In general, a very large dose of rAAV is required in clinical applications, hence manufacturing is anticipated to be carried out in large scale to meet the demand. Current rAAV production relies on multiple plasmid transfections or helper virus infections, which pose challenges in large-scale manufacturing. Additionally, rAAV produced in current manufacturing processes often has large fractions of empty viral particles without the payload. Innovations in viral vector manufacturing are called for to increase vector titers and enhance vector quality.

Our aim is to bring into being a cell line, which has all the essential genetic elements for the synthesis of rAAV vectors, that can be induced to generate the recombinant virus. Since the production does not require any plasmid transfection or virus infection, the process is simple and scalable. Additionally, through tuning of the induction conditions, we can control the expression dynamics of different viral components and modulate the full-to-empty particle ratio of the vector to enhance the product quality.

We have taken a synthetic biology approach to establish rAAV producing cell lines by replacing the native transcriptional regulation of viral genes with inducible promoters, organizing the essential viral genes into genetic modules, and integrating those modules into the HEK293 cell genome. The genome module contained a rAAV genome encoding GFP; the replication module encodes the Rep68 protein and helper proteins of adenovirus; and the packaging module provided capsid proteins and the Rep52 protein. After

transfection, cells with these three modules integrated into the host genome were screened for rAAV productivity. To facilitate screening, we also developed an assay cell line, which harbors *rep* and helper genes but has no viral genome nor *cap* gene, for titration of infectious rAAV. We successfully isolated producers that can generate infectious rAAV upon induction. By optimizing the inducer ratio and induction timing, better titers with high full particle content were achieved.

Systems biology approaches, such as transcriptomics and targeted proteomics, were adopted to quantify both cellular and viral transcripts and proteins. Time-course profiles of viral transcripts and proteins revealed different expression dynamics of each cell line and their possible influence on rAAV titers and quality. Inhibitory host responses attenuating rAAV production were also unveiled through multi-omics analysis. These understandings led to further modification of vector design and cultural conditions to enhance productivity. We redesigned a new inducible genome module, which effectively repressed wasteful GFP expression. Later, we introduced more copies of the *cap* gene into the cell genome. As a result of these strategic modifications, we successfully developed high rAAV producers capable of producing more than 10^5 vector genome per cell with a percentage of full particles exceeding 40%. Further adaptation of these producers into suspension culture demonstrated their scalability.

Taken together, utilizing a systems design and synthetic biology approach, a scalable and controllable rAAV production system was demonstrated. This work exemplifies the crucial role of synthetic biology to engineer next-generation gene therapy manufacturing platforms.

Table of Contents

Abstract	iii
List of Tables	x
List of Figures	xi
List of Abbreviations	xiv
1 Introduction.....	1
1.1 Thesis Organization	3
2 Construction of an rAAV Producer Cell Line through Synthetic Biology	5
2.1 Summary	5
2.2 Introduction.....	6
2.3 Materials and Methods.....	10
2.3.1 Vector Design	10
2.3.2 Generation of rAAV Producer Cell Lines	11
2.3.3 Cell Culture and rAAV Vector Production.....	12
2.3.4 rAAV Preparation and Characterization	12
2.3.5 Illumina Sequencing	13
2.3.6 Targeted Quantitative Proteomics.....	14
2.4 Results and Discussion	16
2.4.1 Motivation and Strategies	16

2.4.2	Vector Design Framework	16
2.4.3	Inducible rAAV Producer Cell Line and Assay Cell Line	19
2.4.4	Capsid-Enhanced Producer Cell Line	24
2.4.5	Inducible Packaging Cell Line	25
2.4.6	Targeted Proteomics Assay for Absolute Protein Quantification	29
2.4.7	Tunable Induction and rAAV Titer	33
2.4.8	Tuning Product Quality and Productivity	36
2.5	Conclusion	41
3	Enhancing the Production of Recombinant Adeno-Associated Virus in Synthetic Cell Lines through Systematic Characterization	42
3.1	Summary	42
3.2	Introduction	43
3.3	Materials and Methods	47
3.3.1	Vectors	47
3.3.2	Generation of rAAV Producer Cell Lines	47
3.3.3	Cell Culture and rAAV Vector Production	47
3.3.4	rAAV Preparation and Characterization	48
3.3.5	Targeted Quantitative Proteomics	48
3.3.6	Transmission Electron Microscopy and Cryo-EM	48
3.3.7	Western Blot	50

3.3.8	Flow Cytometry Assay	50
3.4	Results.....	50
3.4.1	Controlled Gene Expression of GOI Increased Productivity.....	50
3.4.2	Dynamics of Viral Gene Expression and Virus Assembly	56
3.4.3	Cell Line-Dependent Optimal Induction Profile for High Productivity ...	60
3.4.4	Inhibition of Proteasomal Degradation Increased Productivity.....	65
3.4.5	Full Particle Content of rAAV Produced.....	69
3.5	Discussion	71
4	Elevating Recombinant Adeno-Associated Virus Titer and Quality by Modulating Viral Gene Expression Dynamics in High-Producing Synthetic Cell Lines	76
4.1	Summary	76
4.2	Introduction.....	77
4.3	Materials and Methods.....	81
4.3.1	Vector Construction	81
4.3.2	Generation of rAAV Producer Cell Lines	81
4.3.3	Cell Culture and rAAV Production.....	82
4.3.4	rAAV Preparation and Characterization.....	82
4.3.5	Targeted Quantitative Proteomics.....	83
4.4	Results.....	84

4.4.1	Identification of Limiting Factors in Synthetic rAAV Producer Cell Lines	84
4.4.2	Sequential Transfection with The Packaging Module Substantially Increased Capsid and VG titers.....	87
4.4.3	Isolation and Characterization of High-Producing Cell Clones	90
4.4.4	Multiplex Screening for Improving Vector Quality	93
4.4.5	Shifting Balance between Genome Replication and Capsid Formation...	95
4.4.6	Dynamics of Viral Components.....	100
4.5	Discussion.....	101
5	Multi-Omics Kinetic Analysis of Recombinant Adeno-Associated Virus Production by Plasmid Transfection of HEK293 Cells.....	108
5.1	Summary	108
5.2	Introduction.....	109
5.3	Materials and Methods.....	112
5.3.1	Cell Culture and rAAV Production through Triple Transfection	112
5.3.2	rAAV Preparation and Characterization	113
5.3.3	Illumina Sequencing and Differential Expression Analysis	113
5.3.4	Quantitative Proteomics Analysis using Tandem Mass Tags.....	113
5.3.5	Targeted Quantitative Proteomics Analysis.....	114
5.3.6	Functional Class Analysis.....	114

5.4	Results and Discussion	115
5.4.1	Summary of Data	115
5.4.2	Dynamics of Viral Transcripts and Proteins upon Triple Transfection..	119
5.4.3	Host Cell Transcriptome and Proteome Changes Following Triple Transfection	125
5.4.4	Cellular Responses – Integrated View	131
5.4.5	Modulation of Host Responses for Enhancing rAAV Production.....	135
5.5	Concluding Remarks.....	140
6	Summary and Concluding Remarks	141
7	Bibliography	145

List of Tables

Table 2-1. Comparison of transducing titers of a rAAV2-RSS vector provided by ATCC and measured using RM4 assay cells.....	12
Table 2-2. Synthetic heavy isotope-labelled peptides used in targeted proteomics assay.	15
Table 2-3. Summary of identified proteins of interest by LC-MS.....	30
Table 2-4. Candidate peptides belonging to proteins of interest that can be used for quantification.	31
Table 2-5. Selected surrogate peptides based on the preliminary data obtained by PRM.	33
Table 3-1. Synthetic heavy isotope-labeled peptides used in PRM.....	49
Table 4-1. Sequences of heavy isotope-labeled peptides used in the targeted proteomics assay for absolute quantification (AQUA).	83
Table 4-2. Estimation of the number of protein molecules per cell.....	83
Table 5-1. Sequences of heavy isotope-labeled peptides used in the targeted proteomics assay for absolute quantification (AQUA).	114
Table 5-2. Clustering of functional classes identified in different paired comparisons of transcriptome by overrepresentation analysis using DAVID.	126
Table 5-3. Functional class overrepresentation analysis (ORA) using DAVID and gene set enrichment analysis (GSEA) of 48 hpt vs. Mock comparison.....	128
Table 5-4. Clustering of functional classes identified in different paired comparisons of proteome by overrepresentation analysis using DAVID.	130
Table 5-5. Small molecules, and agonists of pathogen pattern recognition signaling used to evaluate the effect of intervention of pathways of interest on rAAV production.....	135

List of Figures

Figure 2-1. AAV genome structure and function of viral proteins.....	8
Figure 2-2. A cell line capable of replicating rAAV genome.....	18
Figure 2-3. Construction of rAAV2 producer cell lines.	22
Figure 2-4. Fluorescence microscope images of RM4 assay cells upon transduction of rAAV2-GFP.....	23
Figure 2-5. Diagram of unspliced and spliced <i>cap</i> transcripts and primers for qPCR analysis.	23
Figure 2-6. Capsid-enhanced rAAV producer cell lines.....	26
Figure 2-7. Inducible rAAV packaging cell lines.....	27
Figure 2-8. Flexibility of inducible rAAV packaging cell lines.	28
Figure 2-9. The workflow of discovery-based proteomics by mass spectrometry.	30
Figure 2-10. The workflow of targeted quantitative proteomics using PRM.....	32
Figure 2-11. Effect of inducer concentrations on viral transcript levels and titers.....	34
Figure 2-12. Tunable viral transcript and protein abundance and product quality.	37
Figure 2-13. Comparison between transfection systems and Pf3 cell line.	38
Figure 2-14. Performance comparison of triple transfection, producer cell line and packaging cell line.	40
Figure 3-1. Construction and characterization of GX rAAV2 producer cell lines.	52
Figure 3-2. Pf3 omics data and cell line development workflow.	54
Figure 3-3. Kinetic profiles of viral transcripts and proteins, total intracellular virus genome, assembled capsids, and vector genome in GX cell lines upon induction.	57

Figure 3-4. Western blot analysis of VP1/2/3 proteins.....	58
Figure 3-5. The specific viral component data of GX cell lines.....	59
Figure 3-6. Screening of induction conditions on the transducing unit titer.	61
Figure 3-7. Induction conditions affected rAAV production by GX6 and GX7 cell lines.	64
Figure 3-8. Screening of proteasome inhibitors and incubation conditions.	66
Figure 3-9. Treatment of proteasome inhibitor MG132 increased viral protein abundance and rAAV titers.....	67
Figure 3-10. The effect of MG132 on viral protein levels was further evaluated by flowcytometry	68
Figure 3-11. Full particle contents produced using GX cell lines under optimal induction conditions with or without MG132 treatment.....	70
Figure 3-12. Electron microscopic examination and quantification of full particles	71
Figure 4-1. Identifying limiting factors in GX rAAV2 producer cell lines.....	86
Figure 4-2. Development of rAAV2 high-producing cell pools by sequential transfection.	88
Figure 4-3. Sorting BFP-positive cells by FACS and evaluating rAAV productivity of SX cell pools by the transduction assay.....	89
Figure 4-4. Single-cell cloning and productivity screening of clones by transduction assay.	91
Figure 4-5. Characterization of cell lines and multiplex screening of induction profiles.	92
Figure 4-6. Cell line stability and suspension adaptation.	93

Figure 4-7. Induction conditions affected rAAV production by GX6A and GX6B cell lines.	96
Figure 4-8. Integrated kinetic profiles of viral elements in GX6A cells.	98
Figure 4-9. Integrated kinetic profiles of viral elements in GX6B cells.....	99
Figure 4-10. Stabilization of Rep68 with Shield1 negatively affected rAAV production.	99
Figure 4-11. Iterative enhancement of synthetic rAAV producer cell lines.	102
Figure 5-1. The scheme of multi-omics analysis of HEK293 cells during rAAV production through plasmid transfection.....	112
Figure 5-2. Principal component analysis (PCA) of samples	116
Figure 5-3. Differential expression analysis of RNA-seq data using EdgeR.	117
Figure 5-4. Differential expression analysis of TMT data using Proteome Discoverer.	118
Figure 5-5. The percentage of transcripts whose encoding proteins were identified in TMT-based proteomics.....	119
Figure 5-6. Viral transcript and protein dynamics.	121
Figure 5-7. Schematics of viral genome and dynamics of helper transcripts.	122
Figure 5-8. Protein dynamics determined by TMT-based proteomics analysis.	124
Figure 5-9. Differentially expressed genes involved in identified functional classes and their fold-changes in log2 scale	129
Figure 5-10. Modulation of host responses for enhancing rAAV production.	137
Figure 5-11. Illustration of modulation of signaling pathways using small molecules..	138
Figure 5-12. Dose-dependent effect.....	139

List of Abbreviations

Abbreviation	Full Name	Abbreviation	Full Name
AAP	Assembly-activating protein	kbp	kilobase pairs
AAV	Adeno-associated virus	LacO	Lactose operon
Ad	Adenovirus	LC	Liquid chromatography
ANOVA	Analysis of variance	LTR	Long terminal repeat
AQUA	Absolute quantification	MAAP	Membrane-associated assembly protein
ATCC	American Type Culture Collection	MFI	Mean fluorescence intensity
BFP	Blue fluorescence protein	MOI	Multiplicity of infection
BHK	Baby hamster kidney	MRN	MRE11/RAD50/NBS1
BV	Baculovirus	mRNA	messenger ribonucleic acid
CAG	CMV early enhancer/chicken β actin	MS	Mass spectrum
cDNA	complementary DNA	NLR	NOD-like receptor
Cap	Capsid protein	ORA	Overrepresentation analysis
Cas9	CRISPR-associated protein 9	ORF	Open reading frame
CDS	Coding sequence	P2A	2A self-cleaving peptides
CMSP	Center for mass spectrometry and proteomics	PCA	Principal component analysis
CMV	Cytomegalovirus	PCR	polymerase chain reaction
CDS	Coding sequence	PEI	Polyethylenimine
CRISPR	Clustered regularly interspaced short palindromic repeats	PM	Packaging module
DD	Destabilization domain	POI	Protein of interest
DEG	Differentially expressed gene	PRM	Parallel reaction monitoring
DNA	Deoxyribonucleic acid	PRR	Pathogen recognition receptor
DNase	Deoxyribonuclease	qPCR	quantitative PCR
DBP	DNA binding protein	qRT-PCR	quantitative real-time PCR
ELISA	Enzyme-linked immunosorbent assay	rAAV	recombinant AAV
ESI	Electrospray ionization	rBV	recombinant BV
FACS	Fluorescence-activated cell sorting	rHSV	recombinant HSV
FDA	Food and drug administration	rtTA3	reverse tetracycline controlled transactivator
FDR	False discovery rate	Rep	Replication protein
GFP	Green fluorescence protein	RLR	RIG-I like receptor
GM	Genome module	RM	Replication module

GMP	Good manufacturing practices	RNA	Ribonucleic acid
GO	Gene ontology	RSV	Rous sarcoma virus
GOI	Gene of interest	ssDNA	Single-strand DNA
GSEA	Gene set enrichment analysis	T2A	2A self-cleaving peptides
hPGK	Human phosphoglycerate kinase	TEM	Transmission electron microscopy
hpi	Hours post-induction	TG	Total intracellular virus genome
hpt	Hours post-transfection	TLR	Toll-like receptor
HSV	Herpes simplex virus	TMT	Tandem mass tags
HSVtk	HSV thymidine kinase	TPM	Transcripts per million
HEK	Human embryonic kidney	TU	Transducing units
IRES	Internal ribosomal entry site	UPR	Unfolded protein response
ITR	Inverted terminal repeat	VG	Vector genomes

1 Introduction

Gene therapies aim at generating long-term expression of therapeutic genes *in vivo*, thereby replacing the defective genes responsible for inherited and monogenic diseases (Anguela & High, 2019; Naso et al., 2017). Gene therapy involves delivering therapeutic genes to targeted tissues for protein expression. This delivery is often carried out by viral vectors (Crystal, 2014; Xiao et al., 1996). Among viral vectors used in gene delivery, adeno-associated virus (AAV) is a superior vehicle due to its nonpathogenic and predominately nonintegrating nature, sustainable transgene expression with a stabilized form (Dunbar et al., 2018), and diverse serotype-dependent tissue tropism (Srivastava, 2016). AAV *in vivo* administration was first successfully applied to a patient to make severe hemophilia B milder in the late 1990s (Kay et al., 2000). Recently, AAV-mediated gene therapy has gained a number of clinical successes. Several products have been approved by Food and Drug Administration (FDA) including Luxturna for retinal dystrophy (Russell et al., 2017), Zolgensma for spinal muscular atrophy (Au et al., 2022), Hemgenix for hemophilia B, Elevidys for Duchenne muscular dystrophy, and Roctavian for hemophilia A. However, large doses of rAAV required in clinical applications place a huge challenge on manufacturing (Au et al., 2022; Paulk, 2020).

Transient transfection of HEK293 cells is one of the most widely used rAAV production method. Genetic materials required for rAAV production including a therapeutic gene of interest (GOI) flanked by inverted terminal repeats (ITRs), AAV *rep/cap* genes, and helper viral genes (i.e., *E2A*, *E4*, *VA RNA*) are delivered to HEK293 cells through plasmid transfection (Xiao et al., 1998). While this method can rapidly

generate various rAAV, the need of plasmid DNA and transfection reagents, and batch-to-batch viability in transfection complex formation pose challenging obstacles in large-scale production (Shupe et al., 2022). Alternatively, rAAV can be produced by helper virus infection such as adenovirus (Chadeuf et al., 2000), herpes simplex virus (Clément et al., 2009), and baculovirus (Joshi et al., 2021). Although these systems offer relatively better scalability, they require good manufacturing practice (GMP)-grade helper viruses and raise safety concerns owing to the potential presence of replication-competent helper viruses and helper viral nucleic acids and proteins (Dobrowsky et al., 2021). Moreover, both transfection and helper virus infection systems generate excessive empty particles, which cause a burden on purification processes and pose immunogenicity issues in clinical applications. To address these challenges, an innovative approach to rAAV manufacturing is called for.

Synthetic biology has been widely employed to engineer and reprogram existing biological systems to exhibit desired behavior or produce therapeutic proteins (Bachhav et al., 2023; Kitada et al., 2018). Advancements in various aspects of synthetic biology, such as DNA synthesis and assembly technology, transgene integration tools, genetic circuits, regulatory elements, and genome editing techniques, have played a crucial role in the success of cellular therapies and the enhancement of biosynthetic processes (El Karoui et al., 2019). For instance, a variety of inducible promoters have been developed that exhibits switch-like behavior, driving and regulating the transcription of genes of interest in response to exogenous signals (e.g., light, antibiotics, hormones, etc.) (Kallunki et al., 2019). Several integration tools have been extensively adopted for the stable integration of transgenes into cell genome to engineer cells for immunotherapies or as cell factories for

therapeutics production. These tools include lentiviral vectors (Oberbek et al., 2011), transposon systems (Balasubramanian et al., 2016; Tipanee et al., 2017), and recombinase-mediated cassette exchange systems (Zhang et al., 2015). Certainly, synthetic biology tools pave the path to the innovation of next-generation medicines and manufacturing platforms.

Systems biology is an extensive, interdisciplinary, and fast-growing field that enables researchers to understand physiology and phenotypes in complex biological systems by integrating information at several molecule levels, from DNA to RNA to protein (Tavassoly et al., 2018). The translation from quantitative data to meaningful biological interpretation was made possible by advanced high-throughput measurement technology, bioinformatics, and computational modeling. Rapid-expanding omics technology, including genomics, transcriptomics, proteomics and metabolomics, has been adopted to enhance growth and productivity of recombinant mammalian cell lines (Lewis et al., 2016). The integration of synthetic biology and systems biology tools allows for the invention of a game-changing manufacturing platform for gene therapy through a virtuous design-build-test-learn cycle.

1.1 Thesis Organization

This thesis represents the application of synthetic and systems biology to establish and enhance a novel rAAV manufacturing platform. Chapter 2 details the construction of a synthetic, inducible rAAV producer cell line by integrating a minimal set of essential viral genes under the control of inducible promoters into the HEK293 cell genome. This cell line can produce rAAV upon small molecule induction and allows for the tuning of full particle content. New methodologies developed for characterization, such as an assay cell line for titration and a targeted proteomics assay for absolute quantification of viral

proteins, are described in detail. Chapter 3 details the enhancement of rAAV production in synthetic producer cells through systematic characterization. It discusses a new vector design capable of repressing wasteful GOI expression, time-course profiles of viral gene expression, and enhanced rAAV productivity through proteasome inhibition. Chapter 4 details the further development of high rAAV producers by identifying a limiting factor and increasing the integrated copies of the limited gene. It provides comprehensive kinetic profiles of viral components and describes enhancement of full particle content through manipulation of inducer concentration and induction time profiles in detail. Chapter 5 details multi-omics analysis of transient transfection of HEK293 cells during rAAV production. It describes cellular responses to rAAV production in detail. Chapter 6 summarizes the key accomplishments of the studies and proposes directions for further enhancement of the technology.

2 Construction of an rAAV Producer Cell Line through Synthetic

Biology

Adapted with permission from Lee, Z.* , Lu, M.* , Irfanullah, E., Soukup, M., & Hu, W. S. (2022) Construction of an rAAV Producer Cell Line through Synthetic Biology. *ACS Synthetic Biology*, 11(10), 3285-3295. <https://doi.org/10.1021/acssynbio.2c00207>
Copyright©2022 by American Chemical Society.

*These authors contributed equally to this work.

2.1 Summary

Recombinant adeno-associated viruses (rAAV) are important gene delivery vehicles for gene therapy applications. Their production relies on plasmid transfection or virus infection of producer cells which pose a challenge in process scale-up. Here we describe a template for a transfection-free, helper virus-free rAAV producer cell line using a synthetic biology approach. Three modules were integrated into HEK293 cells including an rAAV genome and multiple inducible promoters controlling the expression of AAV Rep, Cap, and helper coding sequences. The synthetic cell line generated infectious rAAV vectors upon induction. Independent control over replication and packaging activities allowed for manipulation of the fraction of capsid particles containing viral genomes, affirming the feasibility of tuning gene expression profiles in a synthetic cell line for enhancing the quality of the viral vector produced. The synthetic biology approach for rAAV production presented in this study can be exploited for scalable biomanufacturing.

2.2 Introduction

rAAV vectors are one of the most important gene delivery vehicles for human gene therapy. The many advantages compared to other vectors include tissue tropism by different serotypes, safety and long-term efficacy (Samulski & Muzyczka, 2014). A number of rAAV based products, including Luxturna for retinal dystrophy and Zolgensma for spinal muscular atrophy, have been approved by regulatory agencies for clinical use and generated much excitement for many AAV therapies to come (He et al., 2021). With the new modality comes unique manufacturing challenges distinct from those faced with recombinant protein therapeutics. As more rAAV-based therapies progress through clinical pipelines, there is an ever-increasing need for innovations to make production more scalable and robust.

AAV is a nanoparticle consisting of about sixty capsid proteins of VP1:VP2:VP3 at a 1:1:10 ratio (Gonçalves, 2005) that encapsidates a single-stranded DNA genome of up to 5 kilobases in size (Wu et al., 2010). The AAV genome encodes *rep* and *cap* genes flanked by inverted terminal repeats (ITRs) at the two ends of the genome (Figure 2-1). Viral replication involves the interaction of multiple viral and host cell factors. Both *rep* and *cap* segments are alternatively spliced to generate multiple transcripts and proteins. In addition to rAAV proteins, successful replication requires the presence of helper proteins, typically provided by coinfection by a helper virus such as Adenovirus (Ad). Upon expression of helper genes the large Rep proteins Rep78 and Rep68 in the AAV genome are expressed and nick AAV genomes (Tullis Gregory & Shenk, 2000), facilitating AAV genome replication using the host cell DNA polymerase. The single-stranded DNA-Binding Protein (DBP) encoded in the Ad genome improves processivity of DNA

replication (Ward et al., 1998) and facilitates the formation of sub-nuclear viral replication centers (Hidalgo et al., 2016; Weitzman et al., 1996). The adenoviral E4orf6 protein has many helper functions, including export of viral mRNA from the nucleus, blockage of p53 activity, and stimulation of AAV second-strand synthesis (Leppard, 1997). Adenoviral VA RNAs are non-coding RNAs which bind to and inactivate protein kinase R (PKR), subduing the cellular anti-viral response of protein synthesis arrest (Vachon & Conn, 2016). The *cap* gene encodes VP1, VP2 and VP3 which assemble into capsids. The assembly process is catalyzed by the assembly activating protein (AAP) and localized to the nucleolus (Sonntag et al., 2010). Additionally, the small Rep proteins Rep52 or Rep40 are required for encapsidation of single-stranded AAV genomes into pre-formed capsids (King et al., 2001). More viral factors are continuously being discovered, such as the membrane-associated assembly protein (MAAP) which add further complexity to the viral life cycle (Ogden et al., 2019). The concert of biological reactions complicate efforts to make manufacturing platforms efficient yet consistent.

The most common platform for rAAV vector generation is the transfection of plasmids encoding the therapeutic cassette flanked by AAV ITRs, the AAV *rep* and *cap* genes, and adenovirus helper genes *E2A*, *E4*, and *VA RNA* (Clément & Grieger, 2016). While this method generates vectors readily, the challenges of GMP plasmid preparation, transfection efficiency, and reagent cytotoxicity pose significant hurdles for large scale manufacturing (Srivastava et al., 2021). Many alternative technologies address this issue by integrating some of the necessary components into the host cell genome or helpervirus genome. Stable integration of AAV *rep/cap* and the rAAV genome into HeLa (Chadeuf et al., 2000; Clark et al., 1995) and A549 (Farson et al., 2004) cells resulted in clones which

produced rAAV upon infection with adenovirus. Herpes simplex virus (HSV) was identified as an alternative helpervirus, inspiring a production system where suspension Baby Hamster Kidney cells are co-infected with a recombinant HSV containing the rAAV genome and another recombinant HSV containing AAV *rep/cap* genes (Thomas et al., 2009). In the OneBac system, stable AAV *rep* and *cap* sequences in Sf9 insect cells are induced upon infection with a recombinant baculovirus which harbors the ITR-flanked transgene (Mietzsch et al., 2013). Evidently, these methods still require helper virus infection which come with many operational hurdles such as helpervirus removal in purification (Clément et al., 2009), GMP helpervirus preparation (Wright, 2008) and helpervirus genetic stability (Kohlbrener et al., 2005).

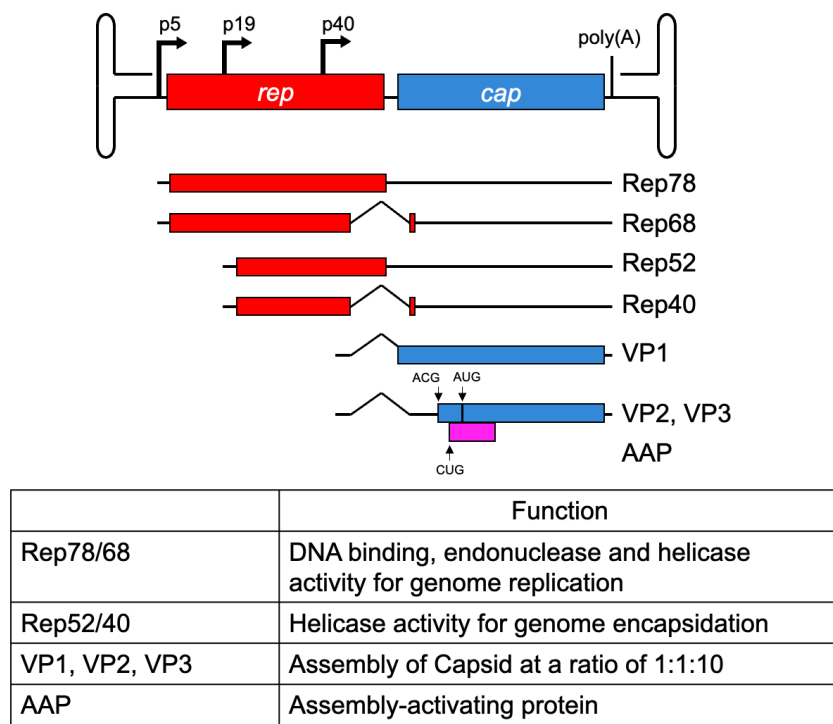


Figure 2-1. AAV genome structure and function of viral proteins.

For decades the main workhorse for protein biotherapeutics production in the biologics industry has been stable producer cell lines which can be cultured in large scale

robustly. Employing stable cell lines for AAV production will eliminate many factors that contribute to process variability and enhance control over product quality. A major barrier to establish a stable AAV producer cell line has been the toxicity of *rep* (Schmidt et al., 2000) and helper genes (Qiao, Li, et al., 2002). To navigate around this challenge effective control of the most toxic protein, Rep78, has been achieved by using inducible promoters including the dexamethasone-inducible mouse mammary tumor virus promoter (Hölscher et al., 1994), the heavy metal-inducible mouse metallothionein promoter (Yang et al., 1994), and the doxycycline-inducible TetON promoter (Chen et al., 2002). Similar promoters have been used to control Adenovirus helper proteins DBP and E4orf6 (Brough et al., 1996; Kovesdi & Hedley, 2010). To control the levels of toxic Rep78/68 and Rep52/40 proteins, a “dual splicing switch” was designed, representing a synthetic intron with a Cre-excisable array of polyadenylation signals (Qiao, Wang, et al., 2002). An early effort to integrate all necessary components for AAV production resulted in cell lines which were not stable (Qiao, Li, et al., 2002). Recently, a fully inducible Rep, Cap, and helper system has been reported, demonstrating the potential of synthetic biology in stable AAV production (Mainwaring et al., 2021).

Being able to produce rAAV is not the only objective for a stable cell line. The resulting platform should generate high quality vectors, especially in the aspect of full capsid content. Patient immune responses pose a great challenge for rAAV therapies (Manno et al., 2006; Mingozzi & High, 2013) and having excess empty capsids will only increase the antigenic load without delivering the therapeutic payload. It has traditionally been the burden of downstream purification efforts to separate the full capsids from empty capsids. Ultracentrifugation methods using CsCl or iodixanol gradients are effective in

separation based on density differences but are not easy to use at large scale (El Andari & Grimm, 2021). Ion exchange chromatography is a more scalable method that can efficiently separate full particles from empty particles based on isoelectric point (Tomono et al., 2018). Nevertheless, low full capsid content and variability between batches put additional burden on downstream processes.

In this study, we constructed stable producer and packaging cell lines of rAAV using synthetic biology tools. We were able to demonstrate control over viral genes and subsequently rAAV titer and full capsid content. This novel strategy enables scalable manufacturing of rAAV with control over vector quality in the production phase.

2.3 Materials and Methods

2.3.1 Vector Design

The rAAV2-GFP genome from pAAV-CAG-GFP (Addgene #37825) was cloned into a Leap-In transposon backbone (ATUM, Inc.) with a puromycin resistance gene driven by a phosphoglycerate kinase promoter, yielding the rAAV genome module (GM). AAV Rep68 sequence was synthesized by Integrated DNA Technologies. The Rep40 start codon ATG was mutated to GGG by PCR, as previously reported (Chejanovsky & Carter, 1989). Rep68 was tagged with mCherry and a mutant FKBP12 destabilization domain (Takara Bio) (Banaszynski et al., 2006). The resulting fusion protein was cloned downstream of the TetON promoter (Takara Bio). The adenovirus E4orf6 and DBP sequences were PCR amplified from the pHelper plasmid (Cell Biolabs, Inc.) and joined by a P2A linker. The resulting CDS was cloned downstream of the GeneSwitch promoter (ThermoFisher Scientific). The Rep68 and helper transcriptional units, as well as the GeneSwitch transactivator gene, were cloned into a Leap-In transposon backbone with a hygromycin

resistance gene linked via T2A to the reverse Tet transactivator, yielding the replication module (RM). AAV *cap* CDS was PCR amplified from pAAV-RC2 (Cell Biolabs, Inc.) and was linked to AAV Rep52 via ECMV IRES and placed under the control of the Cumate Switch promoter (System Biosciences). The transcriptional unit was cloned into a Leap-In transposon backbone with a blasticidin resistance gene and the Cym Repressor, yielding the packaging module C (PM-C). Based on the PM-C, native splice donor sequence was replaced with optimized consensus spliced donor sequence, yielding the PM-D. To build the PM-E, the AAV *cap* CDS was modified to omit the native intron and mutate the VP1 start codon to ACG, as described previously, and termed VP123 (Urabe et al., 2002). The same transcriptional unit of the PM-E was cloned into a Leap-In transposon backbone with a blasticidin resistance gene linked via T2A to the Cym Repressor, yielding the PM-F.

2.3.2 Generation of rAAV Producer Cell Lines

HEK293 cells (Cell Biolabs, Inc.) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% Fetal Bovine Serum (Gibco) and 1X antibiotic-antimycotic (Gibco). Stable cell lines were generated by co-transfection of transposon vectors with Leap-In Transposase[®] mRNA (ATUM, Inc.) at a 1:1 DNA:RNA weight ratio. After three days of outgrowth, cells were passaged into media containing 2 µg/mL puromycin (Invivogen), 200 µg/mL hygromycin B (MilliporeSigma), and/or 10 µg/mL blasticidin (Invivogen). After two weeks of selection, single-cell cloning was performed in 96-wells by limiting dilution. The RM4 assay cell line was developed by integration of the RM into HEK293 cells (Lee et al., 2023). Through enumerating GFP-expressing cells after transduction with samples of rAAV2-RSS (VR-1616), induced RM4 cells gave a comparable transducing titer as the titer provided by ATCC (Table 2-1).

Producer cell lines Pf3 and Pf6 were generated by integration of the GM, RM, and PM-F into HEK293 cells. Packaging cell line RP6 was generated by integration of the PM-F into RM4 assay cells.

Table 2-1. Comparison of transducing titers of a rAAV2-RSS vector provided by ATCC and measured using RM4 assay cells. Titer was determined by flow cytometry after transduction of rAAV2-RSS (8 vg/cell) into induced RM4 cells. Data measured by RM4 represents mean of three replicates.

	rAAV2-RSS (VR-1616) (Data provided by ATCC)	RM4 (Induced)
Transducing units per mL (GFP expressing cells)	5.09E08	3.75E08
95% confidence interval	2.00E08, 9.60E08	3.34E8, 4.16E8

2.3.3 Cell Culture and rAAV Vector Production

For rAAV production via producer cells, cells were seeded at 3.75×10^5 cells per well in 6-well plates with 5 µg/mL doxycycline, 10 nM mifepristone, and 30 µg/mL cumate (Sigma Aldrich), unless noted otherwise. Cells were harvested four days post seeding. For rAAV production via packaging cells, cells were seeded at 7.5×10^5 cells per well in 6-well plates with 5 µg/mL doxycycline, 10nM mifepristone, and 30 µg/mL cumate. One day post seeding, cells were transduced with seed rAAV2-GFP virus, and cells were harvested four days post seeding. Intracellular DNA and RNA were harvested using the Quick-DNA/RNA Miniprep (Zymo Research D7001) and quantified by qPCR and qRT-PCR using SYBR[®] Select Master Mix (ThermoFisher Scientific).

2.3.4 rAAV Preparation and Characterization

Harvested cells were lysed by three cycles of freeze/thaw at -80 °C/37 °C and treated with ≥ 250 units/mL Benzonase (Millipore Sigma E1014) for one hour. Benzonase

activity was quenched by adding 2 mM EDTA, and cell debris was removed by centrifugation. For physical titer assay, proteinase K (Macherey-Nagel, 740506) was added to release DNA, and ssDNA was isolated by column purification (Zymo Research). Quantitative PCR was performed on purified genomes with pAAV-CAG-GFP as a reference standard using primers against the GFP coding sequence to determine the physical titer. Capsid titer was measured using an A20 antibody ELISA kit (Progen PRATV). Transducing titer assays were conducted by seeding RM4 cells at 1.5×10^5 cells per well in 24-well plates with 5 $\mu\text{g/mL}$ doxycycline (Sigma Aldrich) and 10 nM mifepristone (ThermoFisher Scientific). rAAV preparations were added to the cultures one day after seeding, and transducing titers were measured two days post-transduction. The full particle content was obtained by dividing the physical titer by the capsid titer (Grimm et al., 1999).

2.3.5 Illumina Sequencing

Three induction conditions of Pf3 producer cells were investigated: induction with 5 $\mu\text{g/mL}$ doxycycline, 10 nM mifepristone, and 10 $\mu\text{g/mL}$ cumate; induction with 5 $\mu\text{g/mL}$ doxycycline, 10 nM mifepristone, and 90 $\mu\text{g/mL}$ cumate; and no induction. RNA was extracted on day 4 using the Quick-DNA/RNA Miniprep (Zymo Research D7001). Illumina TruSeq Stranded mRNA library preparation was performed by the University of Minnesota Genomics Center. Each library was sequenced as paired-end 150bp reads on a NovaSeq 6000 instrument at a target depth of 20M reads. Reads were mapped using the STAR aligner to the GRCh38 human genome with integrated genome, replication, and packaging module sequences appended. Transcript expression was assessed using Cufflinks and reported as transcripts per million transcripts (TPM).

2.3.6 Targeted Quantitative Proteomics

Cell pellets were reconstituted in 40 μ L of lysis buffer (7.0 M urea, 2.0 M thiourea, 0.4 M Tris, 40 mM chloroacetamide, 10 mM TCEP, 20% acetonitrile). Bradford assay was performed to determine protein concentration. 10 μ g aliquots of each sample were used for proteolytic digestion and brought to the same volume with lysis buffer and incubated at room temperature for 15 min before diluted 4x with water. Heavy isotope-labeled peptides ($K(^{13}C^{15}N)$ or $R(^{13}C^{15}N)$) were purchased from Millipore Sigma and reconstituted in 97.9:2:0.1, H_2O : acetonitrile (ACN): formic acid (FA). Peptide standards were spiked into samples prior to digestion. Trypsin was reconstituted with 10 mM calcium chloride in water and was added to the samples in a 1:40 enzyme to total protein ratio. Samples were incubated 16 h at 37 $^{\circ}C$. Samples were frozen at -80 $^{\circ}C$ and dried in a Speedvac vacuum concentrator after incubation. Once dried, samples were reconstituted in 97.9:2:0.1, H_2O : acetonitrile (ACN): formic acid (FA) and were cleaned up with an MCX-like STAGE tips. The eluted samples were dried in a Speedvac vacuum concentrator.

The dried peptide mixtures were reconstituted in 97.9:2:0.1, H_2O : acetonitrile (ACN): formic acid (FA). 250 ng of each spiked sample were analyzed by capillary LC-MS. A Thermo Scientific UltiMate 3000 RSLCnano system (Thermo Fisher Scientific) with direct column load in 97.9:2:0.1, H_2O :ACN:FA onto a 100 μ m inner diameter 30 cm self-packed LC column was used. Data was acquired on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) in data dependent mode with the following parameters: ESI voltage 2.1 kV, ion transfer tube 275 $^{\circ}C$; Orbitrap MS1 scan 120 k resolution in profile from 380 – 1580 m/z with 50 msec injection time, 100% normalized automatic gain control (AGC); MS2 triggered on the top 15 most abundant ions above 2 x

10⁴ counts with ion trap collision induced dissociation (CID) activation with 35 % collision energy, 10 msec activation time and 0.25 activation Q, 1.6 Da quadrupole isolation window, ion trap detection with 86 msec injection time, 5 x 10⁴ AGC, dynamic exclusion duration 15 sec with +/- 10 ppm mass tolerance. A mass list table with information of peptides of interest was assigned (Table 2-2). Raw files of tandem MS (mass spectra) were analyzed using Skyline (MacLean et al., 2010). The spectra were searched against the database containing sequences of targeted peptides shown in Figure S8B. The following parameters were used to identify peptides of interest: 2 Precursor charges, 1 Ion charges y, b ion types, 0.055 m/z ion match tolerance, 10 ppm product mass tolerance, Label: ¹³C ¹⁵N (C-term, K R), Structural modifications: carbamidomethyl and deamidated. Light-to-heavy peptides ratio were calculated with peak areas, and then ratios were used to determine endogenous protein concentrations.

Table 2-2. Synthetic heavy isotope-labelled peptides used in targeted proteomics assay. The peptide for Cap is common to VP1, 2 and 3. Hence denoted as VP1/2/3.

Protein name	Peptide sequence	Heavy or Light	Targeted m/z (Charge:+2)	Chromatogram
Rep78/68	GIEPTLPNWFAVTK	Light	786.9246	
		Heavy	790.9317	
DBP	NVSLPVAHSDAR	Light	633.3334	
		Heavy	638.3376	
Cap	HPPQILIK	Light	521.8240	
		Heavy	525.8311	
AAP	APTEWVIPR	Light	534.7954	
		Heavy	539.7996	

2.4 Results and Discussion

2.4.1 Motivation and Strategies

We aimed to create a stable cell line which can be induced to produce rAAV vectors without resorting to helper virus infection or plasmid transfection, thereby reducing vector production costs and eliminating the risk of helper virus contamination in the product. Additionally, through separate control of different viral component expression levels we aimed to achieve a higher full capsid content of the produced vector. To accomplish those aims the essential components of making rAAV vectors were grouped into three modules: (1) a genome module containing a rAAV genome including the gene of interest; (2) a replication module providing the large Rep protein of AAV and helper proteins of Ad; (3) a packaging module for expressing capsid proteins and the small Rep protein. Replication and packaging modules used independent inducible promoters and a destabilizing domain that replaced endogenous viral regulation, allowing for controllability of the viral gene expression profile and modulation of the genome content of synthesized rAAV vectors. The expression of Rep and helper proteins elicit cytotoxicity that may lead to long term cell line instability. This concern can be alleviated by using inducible expression and screening for stable clones. Using this design framework and employing transcriptomic and targeted proteomic analysis, the design-build-test cycle can lead to an industrially relevant rAAV production system.

2.4.2 Vector Design Framework

Our design framework called for inducible expression of AAV replication and adenovirus 5 (Ad5) helper proteins which were shown to exert cytotoxic effects (Qiao, Li, et al., 2002; Schmidt et al., 2000). It has been reported that either Rep78 or Rep68 alone

was sufficient to produce infectious AAV (Hölscher et al., 1995). Based on the fact that Rep78 contains an extra cytosolic zinc finger domain compared to Rep68 (Berthet et al., 2005) the Rep68 coding sequence (CDS) was chosen to be cloned into the replication module. The commonly used helper plasmid includes E2A, E4, and VA RNA, but plasmids encoding DBP CDS in place of full-length E2A (Berthet et al., 2005; Matsushita et al., 1998; Xiao et al., 1998) and E4orf6 in place of full-length E4 (Matsushita et al., 1998) have also been used to produce AAV with high titer. We first tested whether cells possessing integrated Rep68 and helper proteins E4orf6 and DBP could be stably maintained in culture and upon induction functionally replicate rAAV genomes. We inserted an ITR-flanked GFP CDS into the AAVS1 site of the HEK293 genome using CRISPR/Cas9 (Figure 2-2A). Subsequently, two lentiviral vectors were used to integrate inducible Rep68 and helper proteins E4orf6 and DBP respectively into the host genome and generate cell line SC5. The capability of SC5 to replicate virus genomes upon induction was indicated by GFP expression (Figure 2-2B) as well as by the several thousand-fold increase in its copy number (Figure 2-2C). While the several thousand-fold increase in viral genome was encouraging, the number needed to increase for a production process.

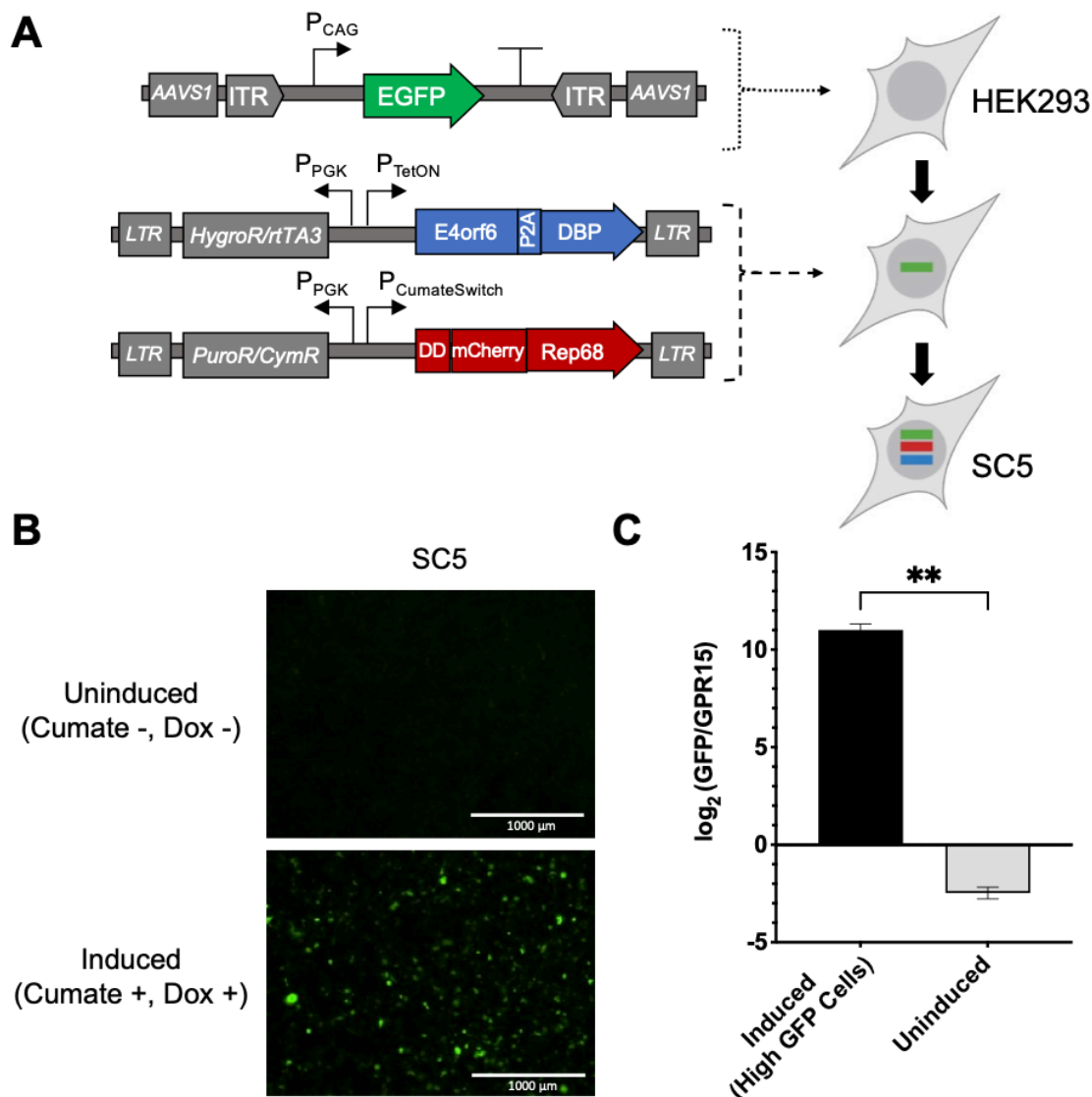


Figure 2-2. A cell line capable of replicating rAAV genome. (A) A rAAV genome module with ITR-flanked GFP CDS was integrated into the AAVS1 site of HEK293 genome using CRISPR-Cas9 editing, and a cell clone was generated. Inducible Rep68, and E4orf6 and DBP proteins in two separate lenti vectors were integrated into the cell clone, yielding the SC5 cell clone. (B) Fluorescence microscopy images of SC5 cell clones induced with 30 μ g/mL Cumate and 5 μ g/mL Dox for six days and the uninduced control. (C) Induced SC5 cells were sorted using fluorescence-activated cell sorting, and virus genome amplification of cells with high GFP expression was determined using qPCR. Data represents mean and standard deviation of triplicate qPCR wells. **: p < 0.01 as determined by a two-tailed, two-sample t-test.

We adopted new vector designs, replacing size-restricted lentiviral vectors with transposon-based vectors for their larger payloads and used vector transfection conditions

which favored multiple copy integration. The idea was that the increased copy number of integrated rAAV genomes would increase genome amplification for rAAV production. We developed a workflow to rapidly build gene cassettes, evaluate infectious rAAV production by transient transfection, then generate stable pools to identify the best designs. Instead of serial transfection as before, triple transfection of three modules and selection was conducted simultaneously to further increase the speed of the screening process and enhance clonal diversity of the resulting pool.

2.4.3 Inducible rAAV Producer Cell Line and Assay Cell Line

We placed the essential components for generation of rAAV into three transposon modules for integration into the host cell genome. The rAAV genome module (GM) contained an AAV-2 ITR flanked GFP CDS (Figure 2-3). The replication module (RM) consisted of (1) Ad5 E4orf6 and DBP linked by P2A under the GeneSwitch promoter control, (2) AAV2 Rep68 tagged by a destabilization domain (DD) and mCherry under the TetON promoter control, and (3) the reverse Tet and GeneSwitch transactivators. In placing Rep68 under TetON control, we omitted the native p5 promoter (Figure 2-1) and removed the internal start codon downstream of p19. The DD is a destabilization domain that tags the protein with an ubiquitination signal for degradation but can be stabilized upon the addition of Shield1 (Banaszynski et al., 2006). The *cap* gene is spliced to form VP subunits and are essential for capsid assembly. The small Rep proteins, Rep52 or Rep40, are essential for AAV genome encapsidation. Both Rep52 and Rep40 have the domain exerting DNA helicase function (King et al., 2001), so only Rep52 and *cap* CDS were included in the packaging module (PM). RNA-Seq data of rAAV production by triple transfection showed that *cap* transcripts were expressed at the highest levels among all

AAV genes (Stutika et al., 2016) (Figure 5-6). Hence, the *cap* gene was placed under the control of the CumateSwitch promoter which conferred high levels of induced expression to replace the native p40 promoter. The RNA-Seq data showed lower abundance of Rep52 transcripts relative to *cap* transcripts. Accordingly, Rep52 was placed downstream of *cap* via an internal ribosome entry site (IRES), which typically results in attenuated translation of the downstream CDS relative to the upstream sequence (Mizuguchi et al., 2000).

HEK293 cells, which constitutively express Ad E1 genes, were transfected with the three modules and ATUM Leap-In Transposase[®] mRNA, and then triple selected with puromycin, hygromycin and blasticidin for 14 days before single cell cloning was conducted to isolate producer cell clones (Figure 2-3A). To quickly evaluate whether producer cell lines could generate infectious rAAV, we used an assay cell line, RM4, which was generated in the same manner as producer cell lines but possessed only the replication module (Lee et al., 2023). The assay cell line would be able to replicate transduced genomes and thus amplify transduction signal, enabling sensitive detection of functional rAAV vectors (Figure 2-4). Upon induction of all three promoters, producer cell clone 2 (PC-2) produced infectious rAAV vectors as evidenced by GFP fluorescence in RM4 assay cells, but the transducing titers were low ($\sim 10^4$ TU/mL) (Figure 2-3B). The low titer may have resulted from deficient virus genome replication and/or capsid assembly.

The *cap* gene encodes three capsid proteins, VP1, VP2 and VP3 which are present in the capsid at 1:1:10 ratio. Primers were designed in exon 2 to quantify total *cap* transcripts and the splice junction to quantify VP2/3 transcripts using qRT-PCR (Figure 2-5). Compared to traditional triple transfection, the total *cap* transcript expression in PC-2 was low, and the level of VP2/3 spliced forms was even lower (Figure 2-3C). These

results suggest that the transcript levels of *cap*, especially the spliced forms were deficient in PC-2.

PC-2 cells were transiently transfected with plasmids encoding supplemental viral genes to assess each's ability to enhance virus production (Figure 2-3D). Two potential enhancing factors were identified, the rAAV genome and the packaging module (PM-C). Higher levels of rAAV genome integration into the host cell genome may enhance successful rescue of virus genomes from the latent, double-stranded form, and more copies of the packaging module may increase transcript and protein levels of *cap* and facilitate capsid assembly. Overexpression of large or small Rep proteins seemed to have an adverse effect on virus production, likely because of their known cytotoxic and cytostatic effects (data not shown) (Qiao, Wang, et al., 2002; Schmidt et al., 2000).

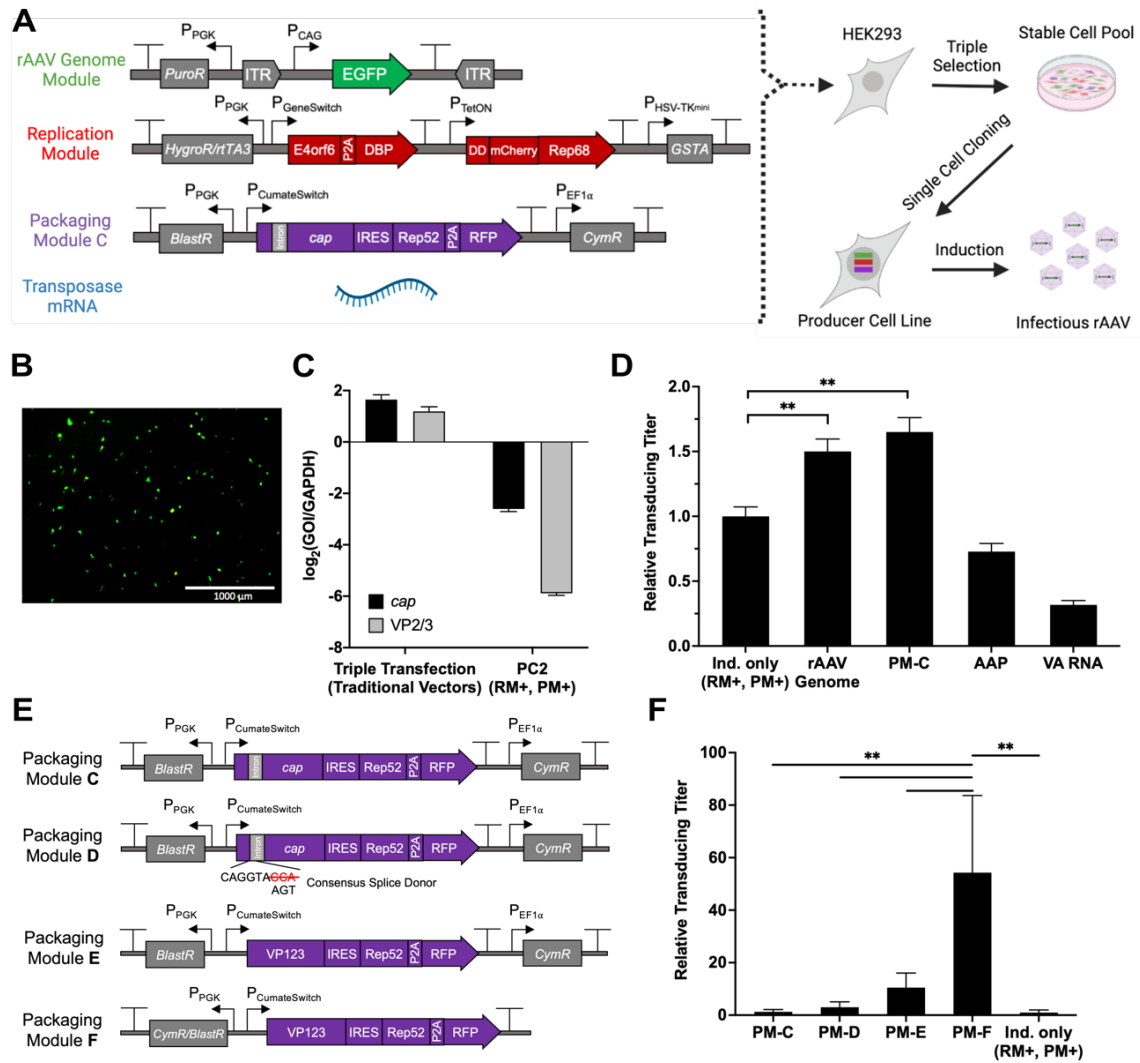


Figure 2-3. Construction of rAAV2 producer cell lines. (A) Gene modules and schematic diagram of rAAV2 producer cell line construction. GSTA denotes the GeneSwitch transactivator gene; rtTA3 denotes the reverse Tet transactivator gene; CymR denotes the Cym Repressor; DD denotes the Destabilization Domain. (B) GFP fluorescence images of RM4 assay cells transduced with rAAV2-GFP produced by producer cell line PC-2. (C) Relative transcript levels of total *cap* and spliced VP2/3 referencing to *GAPDH* in traditional triple transfection of HEK293 cells and induced PC-2 cells. Data represents mean and standard deviation of triplicate qPCR wells. (D) Transducing titers of rAAV obtained through PC-2 cells transiently transfected with viral genes. (E) New packaging modules: control (PM-C), with consensus splice donor (PM-D), with intron-less AAV2 *cap* CDS called VP123 (PM-E), with VP123 and T2A-linked CymR (PM-F). (F) Transducing titers of rAAV obtained by PC-2 cells transiently transfected with new packaging modules. For D and F, the relative titer was in reference to the titer of induction only, and data represents mean and standard deviation of cell counts from multiple images. **: $p < 0.01$ as determined by a two-tailed, two-sample t-test.

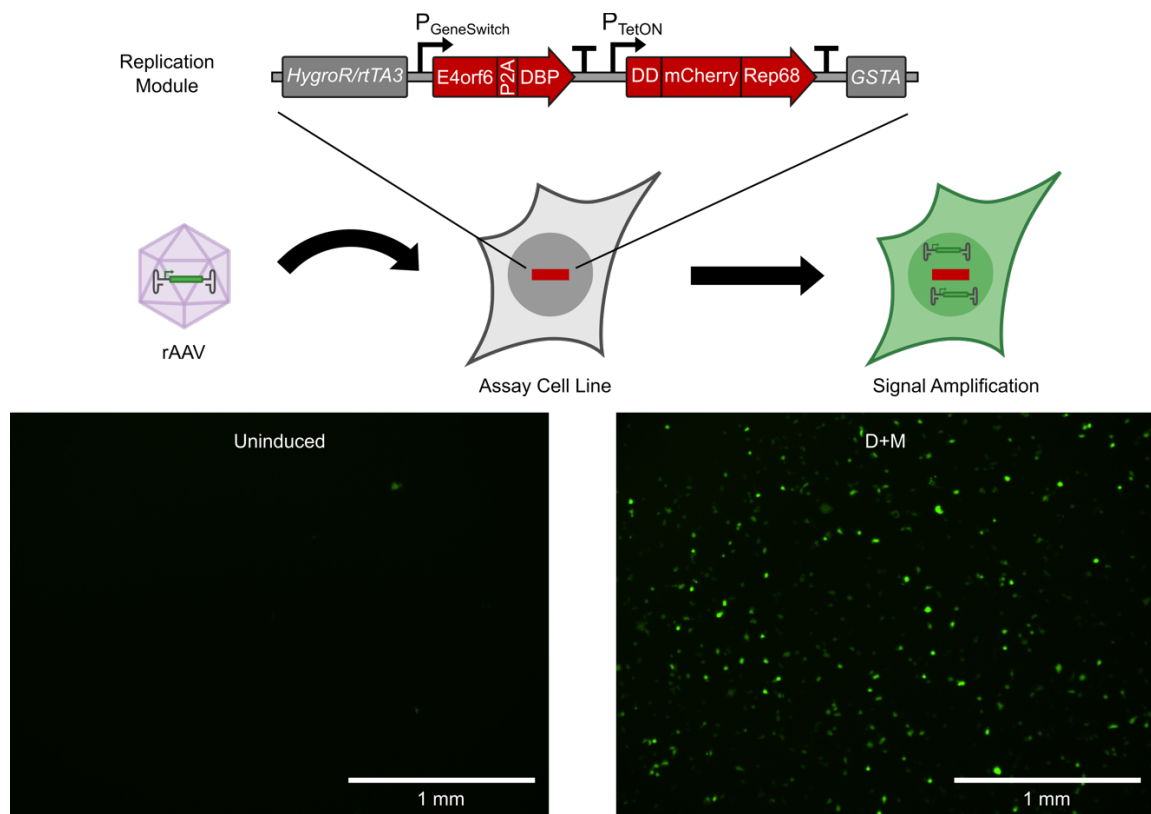


Figure 2-4. Fluorescence microscope images of RM4 assay cells upon transduction of rAAV2-GFP.(A) without induction and (B) with doxycycline and mifepristone induction. A clear increase in sensitivity of detecting transducing events is obtained by inducing Rep68 and helper protein expression in RM4 assay cells.

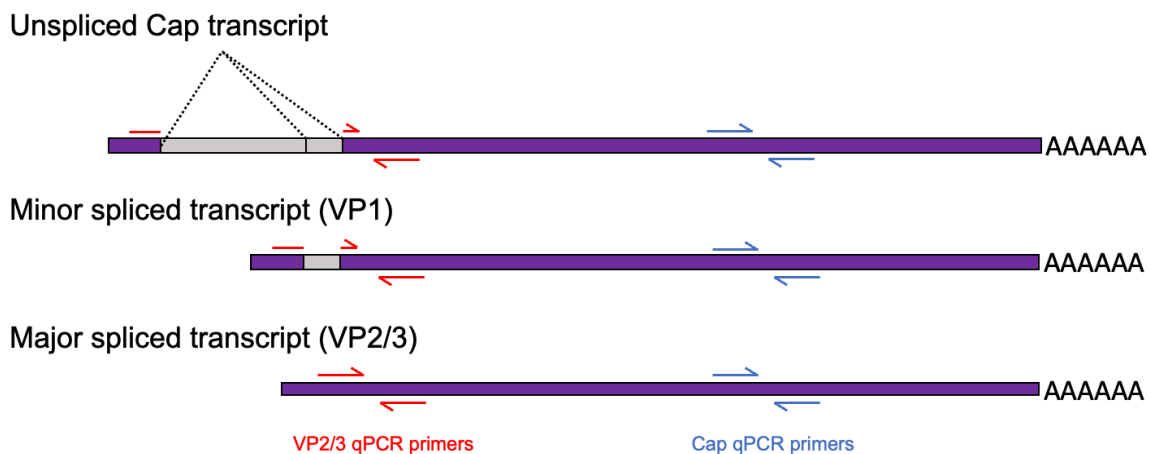


Figure 2-5. Diagram of unspliced and spliced *cap* transcripts and primers for qPCR analysis.

2.4.4 Capsid-Enhanced Producer Cell Line

To increase levels of the spliced transcript encoding VP2/3, three new packaging modules were constructed. One replaced the native splice donor sequence with an optimized consensus splice donor sequence (PM-D) (Figure 2-3E). Another had an intron-less *cap* CDS with translational read-through enabled by an inefficient ACG start codon for VP1, and the resulting CDS was called VP123 (PM-E) (Urabe et al., 2002). The last utilized the same VP123 CDS but consolidated the number of transcriptional units in the transposon by linking CymR and BlastR with T2A (PM-F). PC-2 cells were transiently transfected with these modules and PM-F gave higher virus titer than the others (Figure 2-3F).

HEK293 cells were transfected with the genome module, replication module, and one of the three new packaging modules to generate three cell pools (Figure 2-6A). Consistent with transient transfection, cells with stably integrated PM-F gave substantially higher titer than the others (Figure 2-6B). Two clones isolated from the PM-F pool, Pf3 and Pf6, both produced infectious rAAV vectors as evidenced by GFP fluorescence in RM4 assay cells (Figure 2-6C). Gene copy number analysis showed that all three modules were integrated into the HEK293 genome in multiple copies in both Pf3 and Pf6, except the GFP genome module had relatively lower copy number in Pf6 (Figure 2-6D). Upon induction, Pf3 amplified genome copies to higher levels than Pf6. Before induction, the Rep68 transcript level was very low in both Pf3 and Pf6 (Figure 2-6E). After induction, Rep68 along with DBP and VP123 transcripts were highly expressed, with the VP123 transcript expressed at levels even higher than *GAPDH*. The physical titer, in vector genomes per mL, and the transducing titer, in transducing units per mL, obtained 96 h after induction

were both higher in Pf3 (Figure 2-6F), so Pf3 was used in further experiments. Potential cytotoxicity caused by leaky expression of Rep and helper genes posed concerns about cell stability. Pf3 cells were passaged for about 30 doublings with selection and no significant decrease in physical titer and productivity on a per-cell basis was seen (Figure 2-6G). Considering Pf3 has gone through cell pool and single cell cloning stage before being passaged, the cell line is reasonably stable. We have not yet evaluated the long-term stability and stability in the absence of selective agents, nor have we conducted more extensive cell clone screening for long-term stable high producing cells.

2.4.5 Inducible Packaging Cell Line

Having demonstrated that PM-F had a positive effect on rAAV productivity, we constructed packaging cell lines by integrating RM and PM-F into the host cell genome (Figure 2-7A). Two packaging cell clones, RP6 and RP7, were isolated and both showed induction transcript response profiles very similar to those of producer cell clones (Figure 2-7B). Upon transient transfection of the genome module plasmid under induced conditions, infectious rAAV was produced (Figure 2-7C). To demonstrate the flexibility of a packaging cell line to generate different rAAV vectors by transfecting different genome modules, induced RP6 cells were transfected with four different genome modules to produce various rAAV vectors (Figure 2-8A). rAAV could also be produced by infection of packaging cells with seed rAAV vectors. RP6 infected with seed rAAV2-GFP at different MOI (10, 5, or 1) produced infectious rAAV upon induction (Figure 2-8B). To exclude the possibility that the seed virus used in infection was carried over to the final virus product, a negative control was included in which seed virus was inoculated but RP6 cells were not induced, and very few GFP-positive assay cells were observed.

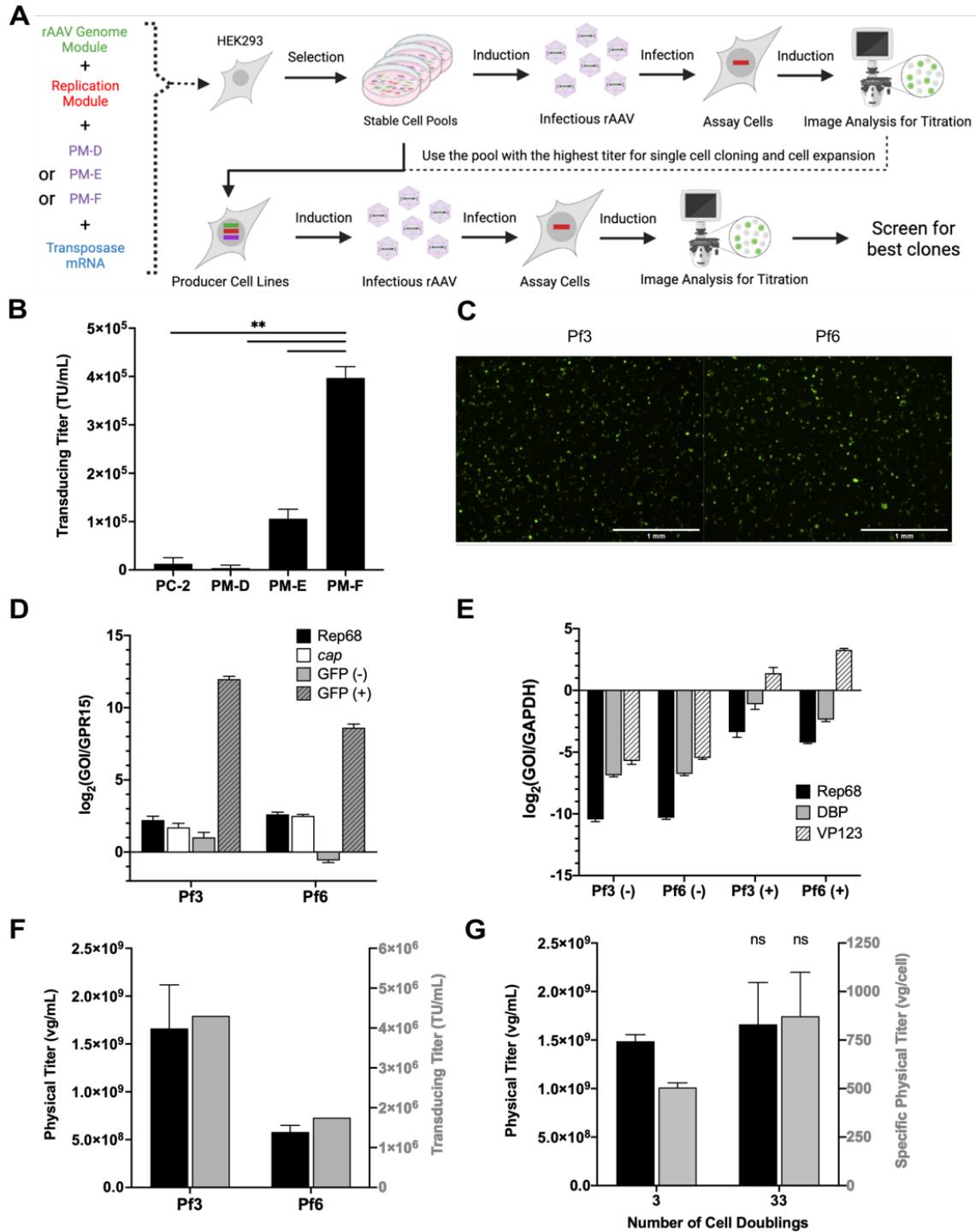


Figure 2-6. Capsid-enhanced rAAV producer cell lines.(A) Schematic diagram of capsid-enhanced rAAV producer cell line construction. (B) Transducing titers of rAAV obtained through the induction of stable cell pools. Data represents mean and standard deviation of cell counts from multiple images. (C) GFP fluorescence images of RM4 assay cells transduced with rAAV2-GFP produced by Pf3 or Pf6 clones. (D) Intracellular copy

numbers of viral genes from each module, and total intracellular virus genomes before induction (-) and 96 hr after induction of PM and RM (+). (E) Relative transcript levels of viral genes in Pf3 and Pf6 cells before induction (-) and at the end of production (+). (F) Virus titers obtained through the induction of Pf3 and Pf6 cells: physical titers measured by qPCR, and transducing titers measured by flow cytometry using RM4 assay cells. (G) Pf3 cell line stability. Physical titer and specific physical titer measured for Pf3 cells induced at 3 and 33 doublings in culture. For D, E and physical titer data, data represents mean and standard deviation of triplicate qPCR wells. **: $p < 0.01$ as determined by a two-tailed, two-sample t-test. ns: $p > 0.05$ as determined by a two-tailed, two-sample t-test.

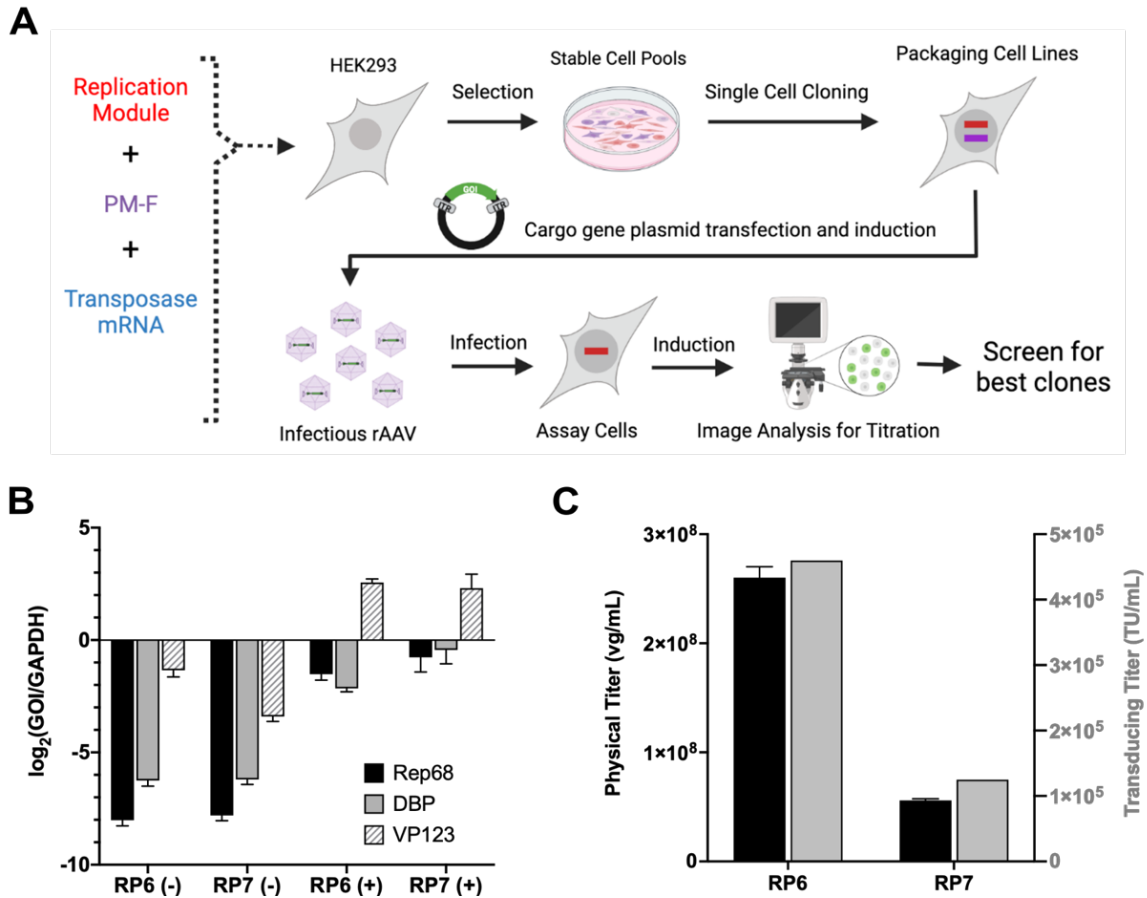


Figure 2-7. Inducible rAAV packaging cell lines. (A) Schematic diagram of a synthetic rAAV2 packaging cell line construction. (B) Relative transcript levels of viral genes in RP6 and RP7 packaging cells before induction (-) and at 96 hr after induction (+). (C) rAAV titers obtained through the transfection and induction of RP6 and RP7 cells: physical titers measured by qPCR, and transducing titers measured by flow cytometry using RM4 assay cells.

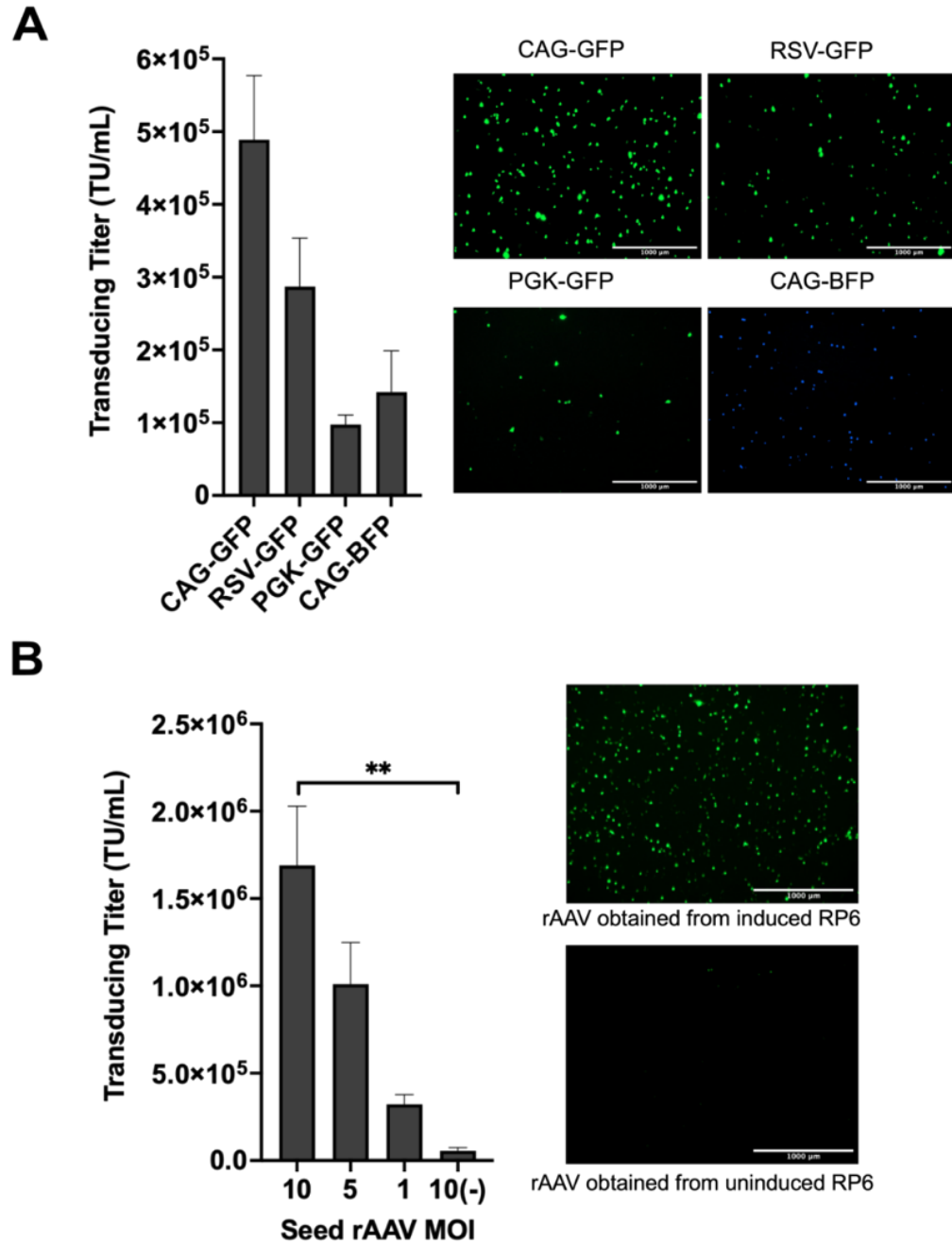


Figure 2-8. Flexibility of inducible rAAV packaging cell lines.(A) Transducing titers of rAAV obtained through RP6 cells transfected with different genome modules and accompanying fluorescence images of transduced assay cells. (B) Transducing titers of rAAV obtained through RP6 cells infected with seed rAAV and accompanying fluorescence images of transduced assay cells. Data represents mean and standard deviation of cell counts from multiple images. **: $p < 0.01$ as determined by a two-tailed, two-sample t-test.

2.4.6 Targeted Proteomics Assay for Absolute Protein Quantification

Genomic and transcriptomic analyses have been widely used to characterize biological systems. However, biological complexity, such as post-transcriptional regulation, can cause disproportionality of protein abundance to messenger RNA levels (mRNA) (Sonenberg & Hinnebusch, 2009). Transcriptomic data might not be directly indicative of protein abundance. Consequently, recent studies tend to use complementary proteomic and transcriptomic analyses to better describe biological systems of interest (Vogel & Marcotte, 2012). To increase understanding about rAAV production, proteomic analysis on transient transfection-based rAAV production via mass spectrometry was first conducted before we employed proteomic analysis to characterize our synthetic cell lines. Transfected HEK293 cells were collected at different timepoints, lysed, and treated by trypsin to digest proteins into oligopeptides for proteomic analysis.

Digested samples were first used for discovery-based proteomics to examine whether proteins of interest (POI) including AAV proteins, Ad helper proteins, and the reporter protein could be identified. The workflow of discovery-based proteomics with Orbitrap Fusion mass spectrometer was described as the following and shown in Figure 2-9. Peptides, as the raw material, were first separated by liquid chromatography (LC) based on differential hydrophobicity or polarity. Then they were ionized into gas phase by electrospray ionization (ESI). The oscillation frequency of ions obtained from Orbitrap was used to calculate the mass-to-charge ratio (m/z). MS spectra were derived from the precursor ions (peptides) and MS/MS spectra were derived from fragment ions (peptide fragments) of a certain precursor ion. Peptide fragmentation generates two series of ions, b- and y- types, that have specific m/z values for peptides sequencing. Peptide-spectrum

match was done computationally with protein databases of human, AAV, Ad and potential contamination to identify peptides and assign identified peptides to specific proteins.

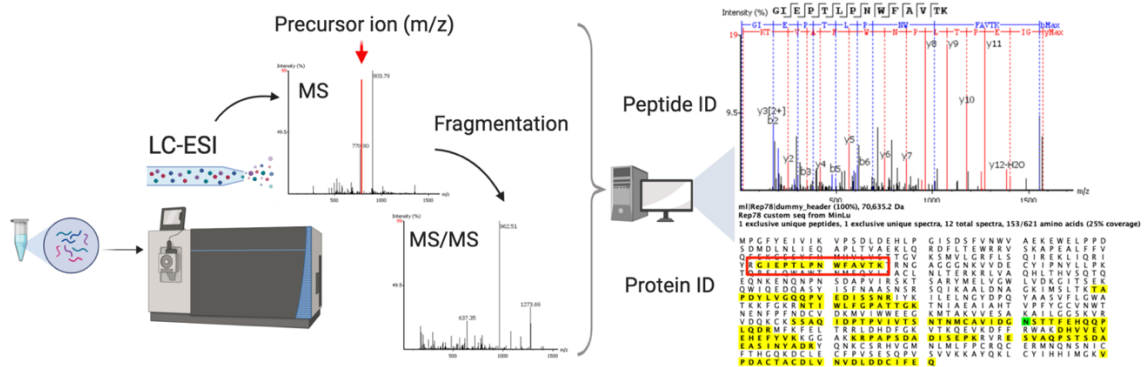


Figure 2-9. The workflow of discovery-based proteomics by mass spectrometry.

As shown in Table 2-3, most of viral POIs were identified, except for some E4 protein isoforms (e.g., E4orf3, E4orf6/7). These undetectable viral proteins may have been present in low abundance. In addition, technical issues such as a high dynamic range of protein content, potential peptide loss during sample preparation, and the method of data acquisition may have also caused these proteins to be undetectable by mass spectrometry.

Another point of interest of this preliminary proteomic analysis was to identify unique peptides that can distinguish known Rep and Cap proteins isoforms and be used for protein quantification. As shown in Table 2-4, the peptides exclusive to viral proteins and different isoforms were detected. The result suggested that the large Rep proteins (Rep78/68) could be distinguished from small Rep proteins (Rep52/40), and VP1 could be distinguished from VP2 and VP3. This information was subsequently used for the development of targeted quantitative proteomics.

Table 2-3. Summary of identified proteins of interest by LC-MS.

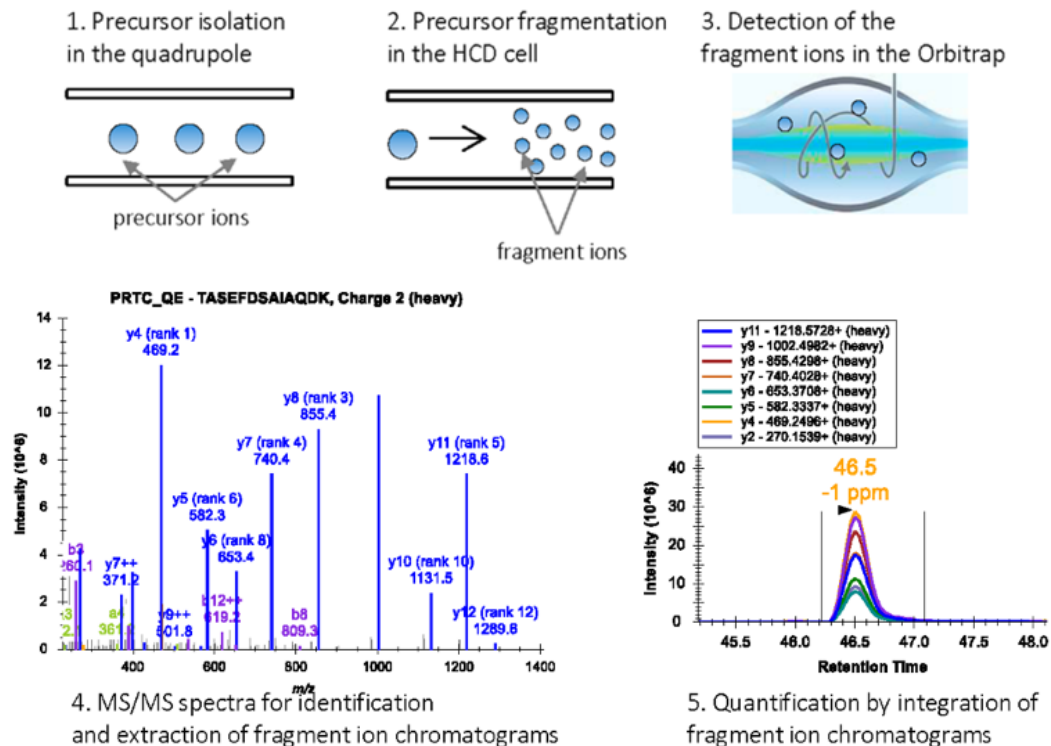
Proteins of Interest	MW	Coverage	Peptide Count
Rep (Rep78/68/52/40)	71 kDa	25%	16
Cap (VP1/2/3)	82 kDa	46%	30
AAP	21 kDa	31%	4
E4orf6	34 kDa	13%	4
DBP	59 kDa	70%	30
GFP	27 kDa	75%	21

Table 2-4. Candidate peptides belonging to proteins of interest that can be used for quantification.

Belonging to	Peptide sequence	Coding region
Rep78/68	GIEPTLPNWF ~~AVTK~~	
	D ~~FLTEWR~~	
Rep	TAPDYLVGQ ~~QPVEDISSNR~~	
	NTIWLFGPATTGK	
	AILGGSK	
VP1	GLVLPGYK	
	QLDSGDN ~~PYLK~~	
VP1/2/3	HPPPQILIK	
	TTNPVATEQYGSVSTNLQR	
AAP	APTEWVIPR	
DBP	NVSLPVAHSDAR	
	LTPFALSNAEDLDADLISDK	
E4orf6	EELVILR	
GFP	FSVSGEGEGDATYGK	
ACTB	AGFAGDDAPR	

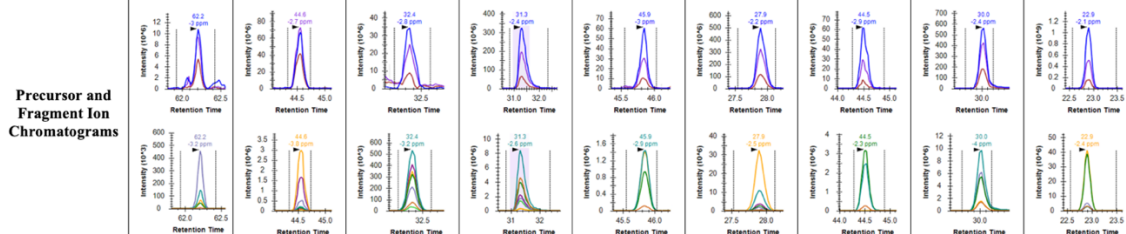
Based on the promising results from discovery-based proteomics, we aimed to quantify identified proteins by targeted quantitative proteomics. The first step was to select suitable surrogate peptides for each POI from all previously detected peptides. Several peptide candidates were analyzed through parallel reaction monitoring (PRM). PRM is a recently developed method for sensitive and specific quantitative proteomics (Peterson et

al., 2012). PRM uses a quadrupole-Orbitrap hybrid mass filter. Quadrupole can specifically select precursor ions of interest to filter out background ions, and Orbitrap provides higher resolution and ability of detecting all potential fragment ions in a single run. These features of PRM allows for both discovery-based and targeted quantitative proteomics (Doerr, 2012). The typical PRM workflow is shown in Figure 2-10. The resulting ion chromatograms can be used for quantification by integrating peak areas of fragment ions (Rauniyar, 2015). Targeted proteomics can quantify several POIs (10s-100s) in a single run of experiment, so it has a higher multiplex power than traditional ELISA (Shi et al., 2016). Incorporated with internal standards like synthetic peptides, absolute quantification (AQUA) across different states of samples can be consistently and reproducibly determined (Picotti & Aebersold, 2012).



The criteria for choosing surrogate peptides include appropriate peptide length, charge states, exclusiveness, less post-translational modification, no miscleavage, and signal intensity shown in PRM data (Rauniyar, 2015). Based on our PRM result, the best surrogate peptides and charge states for each POI were selected and shown in Table 2-5. The high signal intensity and the symmetrical alignment of precursor and fragment ions were the indication for desirable quantotypic peptides. We then synthesized heavy-isotope labeled surrogate peptides that were further used as internal standards for AQUA of viral proteins in our synthetic cells.

Table 2-5. Selected surrogate peptides based on the preliminary data obtained by PRM.

Protein	Rep78/68	Rep	VP1	VP1/2/3	AAP	DBP	E4orf6	GFP	ACTB
Peptide Sequence	GIEPTLPN WFAVTK	TAPDYLVG QQPVEDIS SNR	QLDSGDNP YLK	HPPQILIK	APTEWVIP R	NVSLPVAH SDAR	EELVILR	FSVSGEGE GDATYGK	AGFAGDD APR
Charge State	786.9247++	697.0114+++	625.3006++	521.8240++	534.7955++	633.3335++	436.2660++	752.3335++	488.7278++
Precursor and Fragment Ion Chromatograms									

2.4.7 Tunable Induction and rAAV Titer

One of the key features of our rAAV producing cell design was the controllability of viral gene expression, hence we examined the gene expression responses and rAAV titers under different induction conditions. TetON and CumateSwitch promoters exerted wide ranges of transcriptional control on Rep68 and VP123, respectively, and increasing induction strength increased titers (Figure 2-11A-D). In contrast, the GeneSwitch promoter reached a transcriptional plateau at a relatively low concentration, so it did not show a wide

range of dynamics in the DBP transcript level or rAAV titer (Figure 2-11E, F). To determine if the low Rep68 levels in Pf3 was constraining genome replication, Shield1 was added to stabilize Rep68 upon induction of Pf3. Unexpectedly, the rAAV titer reduced (data not shown). Hence, Shield1 stabilization of Rep68 was not deployed in the studies of rAAV production using Pf3 cells.

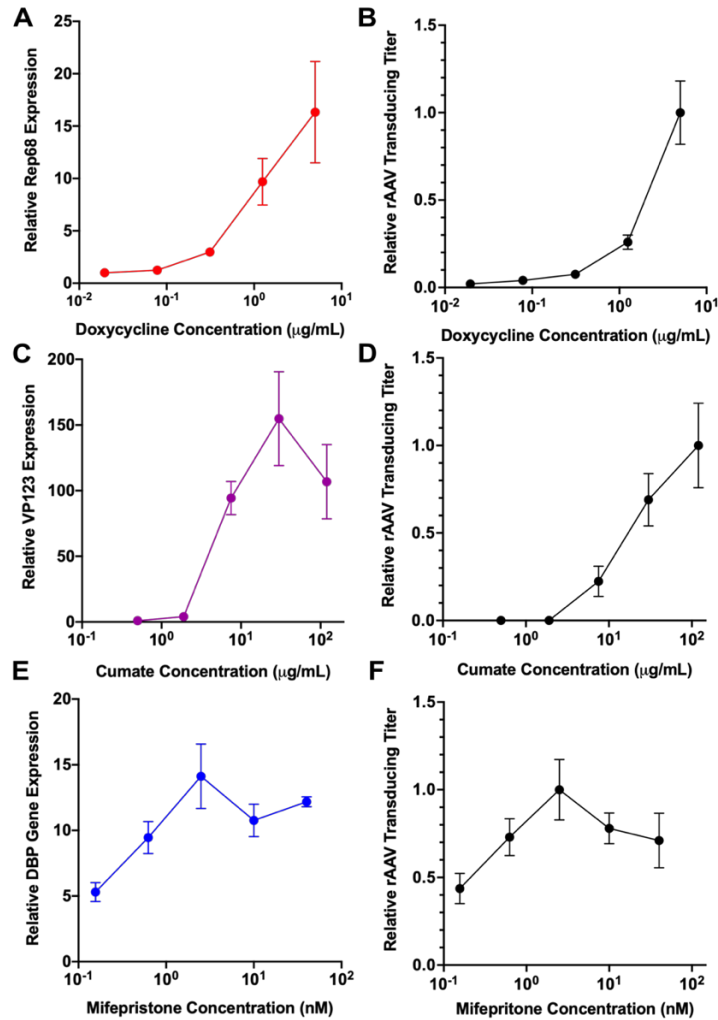


Figure 2-11. Effect of inducer concentrations on viral transcript levels and titers. (A) Effect of doxycycline concentration on Rep68 transcript level, measured by qPCR. (B) Effect of doxycycline concentration on transducing rAAV titer, measured by RM4 assay cells. The relative titer was in reference to the titer obtained with the highest doxycycline concentration. Mifepristone and cumate concentrations were kept at 10 nM and 30 μg/mL, respectively. (C) Effect of cumate concentration on VP123 transcript level, measured by

qPCR. (D) Effect of cumate concentration on transducing rAAV titer, measured by RM4 assay cells. The relative titer was in reference to the titer obtained with the highest cumate concentration. Doxycycline and mifepristone concentrations were kept at 5 µg/mL and 10 nM, respectively. (E) Effect of mifepristone concentration on DBP transcript level, measured by qPCR. (F) Effect of mifepristone concentration on transducing rAAV titer, measured by RM4 assay cells. The relative titer was in reference to the highest titer obtained with 2.5 nM of mifepristone. Doxycycline and cumate concentrations were kept at 5 µg/mL and 30 µg/mL, respectively. qPCR data represents mean and standard deviation of triplicate qPCR wells. Transducing titer data represents mean and standard deviation of cell counts from multiple images.

Considering VP proteins are among the most abundant viral proteins in traditional triple transfection and their possible excess may contribute to high levels of empty viral particles, we evaluated the effect of varying cumate induction concentration on viral transcript and protein expression through RNA-seq and targeted proteomics. Quantitative measurement of protein abundance levels was achieved by spiking into the samples synthetic isotope-labelled unique peptides of Rep68, DBP, VP1/2/3, and AAP. Induction concentrations of doxycycline and mifepristone were kept at 10 µg/ml and at 2.5 nM, respectively. When Pf3 cells were uninduced, Rep68 transcripts were at very low levels while DBP and VP2/3 were more abundant (Figure 2-12A). All three viral transcripts increased to levels similar to *GAPDH*, one of the top ~30 cellular transcripts, or higher upon induction. Increasing the cumate level from 10 µg/mL to 90 µg/mL increased VP2/3 transcripts by roughly 2-fold (Figure 2-12A).

At the protein level, Rep68 and AAP were almost undetectable before induction. Such low transcript and protein expression may have contributed to the absence of cytotoxicity and allowed for clonal growth and eventual isolation. Upon induction, abundance levels of the four viral proteins also drastically increased (Figure 2-12B). Increasing cumate concentration from 10 µg/mL to 90 µg/mL increased VP1/2/3 proteins

and AAP by roughly 2-fold. We estimated the number of protein molecules in cells by assuming a cell contained 360 pg of total protein. DBP had the highest abundance of approximately 10^6 copies per cell, Rep68 and VP proteins reached approximately 10^5 copies per cell, and AAP showed the least abundance of 10^4 copies per cell. Order of magnitude-wise this gives about 10^4 viral particles per cell if all VP proteins were assembled. According to Figure 2-6D, induced Pf3 cells could accumulate 5×10^4 intracellular effective genome copies per cell, thus giving ~1-2 Rep68 proteins per viral genome and ~1 AAP/particle. The proteomic analysis validated viral replication at the protein level, especially for AAP which was undistinguishable from VP2/3 at the RNA level. It also provided an estimate of the molar ratios of different viral components for further cell line optimization.

2.4.8 Tuning Product Quality and Productivity

By varying inducer levels and altering the ratio of various viral components, the properties of the viral particles produced were changed. High cumate induction gave high particle yield, while full dox induction favored high genome amplification. While high cumate levels gave higher particle titer, a smaller proportion of those particles contained genome. Conversely, at a lower cumate concentration and fully induced Dox level, the low particle titer was compensated by a higher level (73%) of full capsids (Figure 2-12C). Using the traditional triple transfection method, we obtained approximately 35% full capsid content. Others have reported full capsid content from below 20% to approximately 50% (Penaud-Budloo et al., 2018). The full capsid content is highly variable, but through the design of genetic modules, the engineered producer cell line Pf3 was able to enhance this quality attribute of its products.

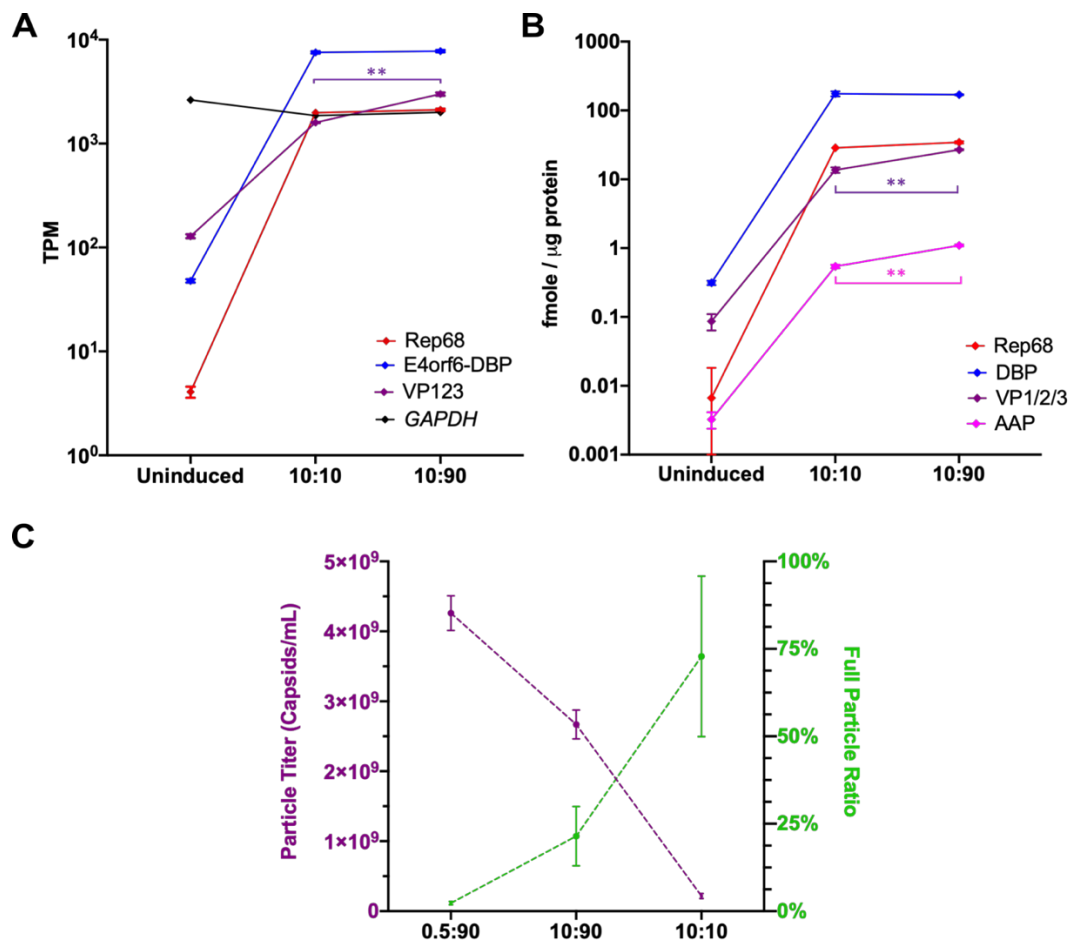


Figure 2-12. Tunable viral transcript and protein abundance and product quality.(A) Transcripts per million of viral genes and GAPDH expressed in Pf3 cells at 96 hr post-induction measured by RNA-sequencing. (B) Absolute quantification of viral proteins expressed in Pf3 cells at 96 hr post-induction measured by parallel reaction monitoring (PRM). (C) Effect of varying dox:cumate ratios on particle titer and full capsid content. Presented numbers for dox and cumate represent the actual concentration of the inducers used, in $\mu\text{g/mL}$, and the mifepristone concentration was kept at 2.5 nM. Data represents mean and standard deviation of three independent replicates. **: $p < 0.01$ as determined by a two-tailed, two-sample t-test.

The viral vector yield of Pf3 was compared to those of triple transfections using the traditional three plasmids and the three modular transposon vectors (Figure 2-13A). Traditional triple transfection gave a higher titer than transposon-module triple transfection, which in turn was significantly higher than Pf3. At the transcript levels the traditional triple transfection method had significantly higher levels of Rep68 and DBP than both Pf3 and

transposon-module triple transfection, but not for VP2/3 (Figure 2-13B). Transposon-module triple transfection gave a higher titer than Pf3, but there was no significant difference in transcript levels of Rep68, DBP and VP2/3.

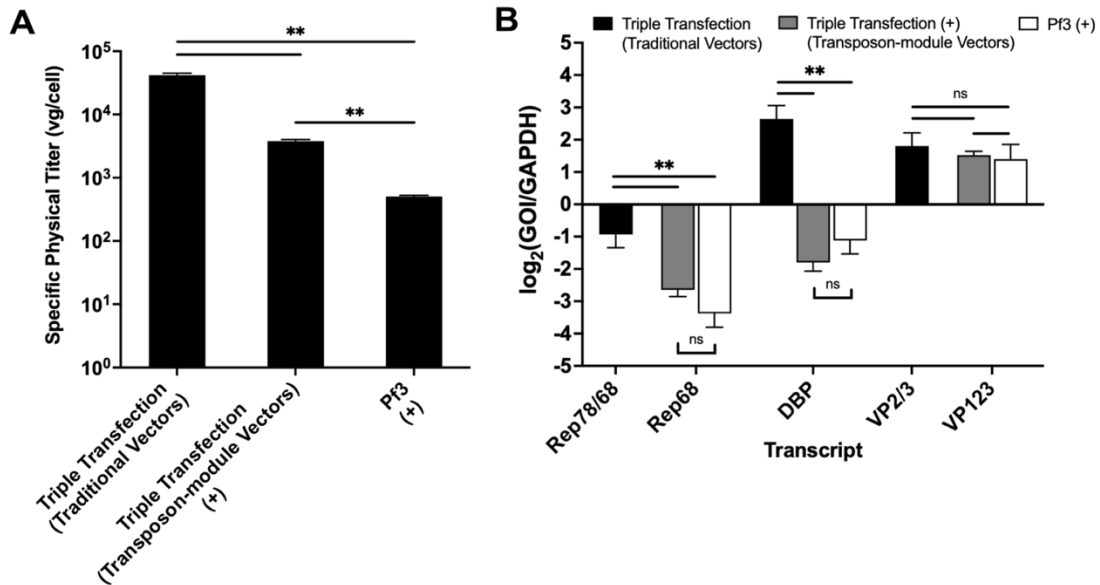


Figure 2-13. Comparison between transfection systems and Pf3 cell line. **(A)** Specific physical titers of rAAV produced from traditional triple transfection, transposon-module triple transfection, and Pf3 production methods. **(B)** Relative transcript levels of viral genes at the end of production of the three production methods. Data represents mean and standard deviation of triplicate qPCR wells. **: $p < 0.01$ as determined by a two-tailed, two-sample t-test. ns: $p > 0.05$ as determined by a two-tailed, two-sample t-test.

Although the production using different methods reported in this study was mostly carried out in dishes and was hardly optimized, we plotted their performance data together for a comparison (Figure 2-14). The triple transfection-based methods gave much higher titers than either Pf3 or RP6. Possible causes of differences in titers among the three methods are many. Understanding factors causing those differences may help enhance the productivity and product quality. Triple transfection using traditional viral gene-based vectors gave about ten-fold higher titer than the transposon-module vectors. The large sizes of the three transposon-module vectors (9,12, and 15 kbp) likely led to a lower transfection

efficiency compared to the smaller traditional plasmids (5, 7, and 11 kbp). The expression efficiency using native viral regulatory systems as in the traditional transfection might be higher than using the heterologous promoters. In addition, the significant difference in DBP transcript levels could potentially be a limiting factor (Figure 2-13B). The kinetics of viral protein synthesis and viral genome amplification are different between plasmid transfection and induction of transcription. Plasmid transfection leads to a rapid increase of viral proteins and viral genomes, whereas transcription induction would result in a more gradual increase in viral protein abundance and genome replication. Further exploiting the versatility of our inducible constructs to exert optimal temporal control of gene expression is warranted.

The productivity of Pf3 is still two orders of magnitude lower than the traditional triple transfection-based production, although the virus produced had comparable VG/TU ratio (Figure 2-14D). It is worth noting that such a synthetic virus production cell line eliminates the hurdle of plasmid production and transfection, and thus enables a scalable manufacturing process. Intensified cell culture systems through perfusion culture employing HEK293 cells or fed-batch culture using insect cells has long been practiced and awaits exploitation (Grieger et al., 2016; Joshi et al., 2019; Kyung et al., 1994). Further work to optimize the design of the genetic modules, modulate host cell's anti-viral responses, and screening for long term stable high producers will likely increase the productivity by at least an order of magnitude and make large scale production of rAAV a reality.

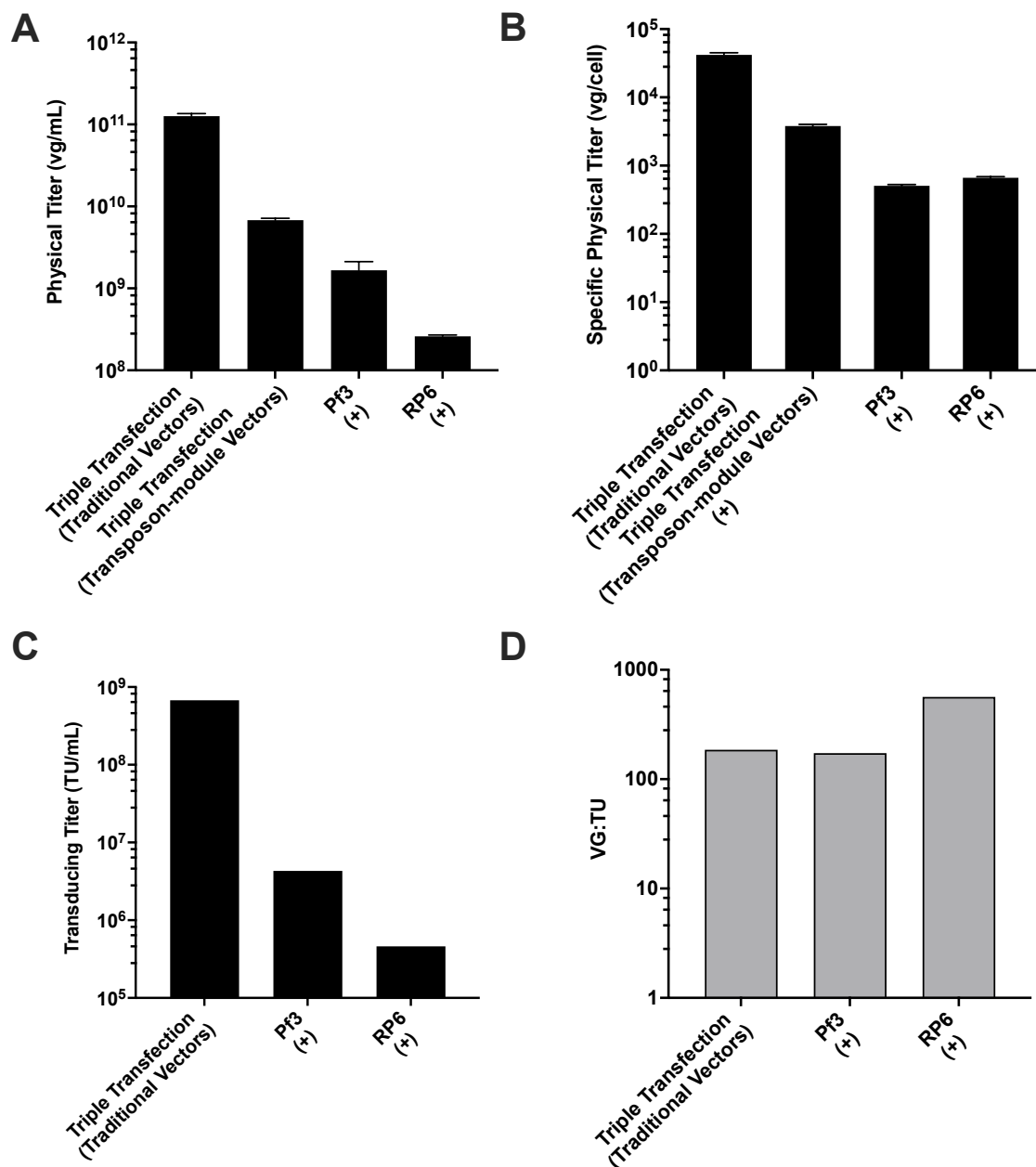


Figure 2-14. Performance comparison of triple transfection, producer cell line and packaging cell line. (A) Virus titers obtained through triple transfection, induction of Pf3 cells, and transfection and induction of RP6 cells: Physical titers measured by qPCR. (B) Specific physical titers of rAAV produced from triple transfection, Pf3 and RP6 cells. (C) Transducing titers measured by flow cytometry using RM4 assay cells. (D) Physical titer to transducing titer ratio. Physical titer data represents mean and standard deviation of triplicate qPCR wells.

2.5 Conclusion

In conclusion, we have taken a synthetic biology approach to design and construct a cell line capable of fully inducible production of rAAV particles. Low basal transcription and high induced transcription of viral genes and genome amplification were achieved. Independent control over replication and packaging modules enabled tuning of the vector-containing full particle content. By leaving out the genome module, we also demonstrated a packaging cell line which can be used to produce a variety of rAAV upon transfection of different genome module vectors. While we used AAV2 capsid proteins in both production and packaging cell lines, the strategy can be readily adopted to generate cell lines for producing or packaging rAAV of other serotypes. Stable cell lines are key to establish scalable rAAV manufacturing. The present work exemplifies the crucial role of synthetic biology to engineer next-generation gene therapy manufacturing platforms.

3 Enhancing the Production of Recombinant Adeno-Associated Virus in Synthetic Cell Lines through Systematic Characterization

Reproduced with permission from Lu, M., Lee, Z., Lin, Y. C., Irfanullah, I., Cai, W., & Hu, W. S. (2023). Enhancing the Production of Recombinant Adeno-Associated Virus in Synthetic Cell Lines through Systematic Characterization. *Biotechnology and Bioengineering*, 1-14. <https://doi.org/10.1002/bit.28562>

3.1 Summary

Recombinant adeno-associated virus (rAAV) is among the most commonly used *in vivo* gene delivery vehicles and has seen a number of successes in clinical application. Current manufacturing processes of rAAV employ multiple plasmid transfection or rely on virus infection and face challenges in scale-up. A synthetic biology approach was taken to generate stable cell lines with integrated genetic modules, which produced rAAV upon induction albeit at a low productivity. To identify potential factors that restrained the productivity, we systematically characterized virus production kinetics through targeted quantitative proteomics and various physical assays of viral components. We demonstrated that reducing the excessive expression of gene of interest by its conditional expression greatly increased the productivity of these synthetic cell lines. Further enhancement was gained by optimizing induction profiles and alleviating proteasomal degradation of viral capsid protein by the addition of proteasome inhibitors. Altogether, these enhancements brought the productivity close to traditional multiple plasmid transfection. The rAAV produced had comparable full particle contents as those produced by conventional transient plasmid transfection. The present work exemplified the versatility of our synthetic biology-

based viral vector production platform and its potential for plasmid- and virus-free rAAV manufacturing.

3.2 Introduction

Recombinant adeno-associated virus (rAAV) is one of the most commonly used delivery vehicles for gene therapy. Its many advantages include effective *in vivo* gene delivery, long-term persistency, different tissue tropism by various serotypes, and high safety (Samulski & Muzyczka, 2014). Several rAAV-based products, including Luxturna for retinal dystrophy, Zolgensma for spinal muscular atrophy, and Hemgenix for hemophilia B have been approved by the United States Food and Drug Administration in recent years, paving the way for more rAAV-based therapies to come (Herzog et al., 2023; Li & Samulski, 2020). However, a very large dose of rAAV is required in clinical applications (Mendell et al., 2017), and the manufacturing of large quantities of rAAV poses a significant challenge. Innovations in rAAV vector manufacturing are called for to create a scalable and robust production platform to meet the demand.

AAV is a small, nonenveloped capsid virus with a single-stranded DNA genome encoding two genes, *rep* and *cap*, flanked by inverted terminal repeats (ITRs) (Sonntag et al., 2010; Srivastava et al., 1983). AAV as a replication-defective virus, needs the coinfection of a helper virus such as Adenovirus (Ad) to have efficient *rep/cap* gene expression and virus genome replication (Stutika et al., 2016). With alternative splicing and multiple translation initiation sites, both *rep* and *cap* genes generate multiple transcripts and proteins. The large Rep proteins, Rep78 and Rep68, have multiple functions, including DNA binding, ATP-dependent endonuclease, and helicase activities (Im & Muzyczka, 1990) to bind and nick ITRs for AAV genome replication. The small Rep

proteins, Rep52 and Rep40, use their helicase activity to encapsidate AAV genomes into pre-formed capsids (King et al., 2001). The *cap* gene encodes VP1, VP2, and VP3 capsid subunits which at a 1:1:10 ratio assemble into the 60-subunit capsid. After translation in the cytoplasm, their translocation into the nucleolus and assembly is facilitated by the assembly-activating protein (AAP) (Sonntag et al., 2010). Recently, the membrane-associated assembly protein (MAAP) expressed from a novel open reading frame (ORF) in the *cap* gene was found to modulate secretion of AAV (Elmore et al., 2021; Ogden et al., 2019). Successful AAV genome replication also requires the participation of components from Ad, such as DNA-binding protein (DBP) which increases the processivity of DNA replication (Ward et al., 1998) and E1B55K/E4orf6 complex, which mediates degradation of host cellular proteins such as the MRN complex and p53 (Querido et al., 2001; Schwartz et al., 2007).

To make AAV a gene therapy agent, the *rep* and *cap* genes in the genome are replaced by a gene of interest (GOI) which is to be expressed upon delivery to target cells. The resulting rAAV genome thus lacks the two essential genes for its replication. rAAV is most commonly produced by triple transfection of HEK293 cells with three plasmids, one containing the rAAV genome and two others providing *rep/cap* genes and helper viral genes including *E2A*, *E4* and *VA RNA* in trans, respectively (Matsushita et al., 1998). Additional Ad helper genes, *E1A* and *E1B*, are already part of the HEK293 genome (Lin et al., 2014). This method produces rAAV readily but is cumbersome in large-scale manufacturing as large quantities of multiple plasmids must first be produced (Srivastava et al., 2021).

Alternative methods of rAAV production use recombinant helper viruses, such as co-infecting Baby Hamster Kidney (BHK) cells with two strains of replication-defective recombinant herpes simplex viruses (rHSV) containing rAAV genome and *rep/cap* genes, respectively (Clément et al., 2009). A similar approach using recombinant Ad also produced rAAV (Merten et al., 2005). The potential of contamination by helper viruses in rAAV products requires additional downstream purification (Penaud-Budloo et al., 2018). To alleviate the concern about residual adenovirus contamination in the product, a tetracycline-enabled self-attenuating Ad was used (Su et al., 2022). Production systems using native *rep/cap* genes present the possibility of generating replication-competent AAV (rcAAV) through homology between the ITR and p5 promoter and through nonhomologous recombination (Allen et al., 1997). rAAV was also produced in a baculovirus-based platform by coinfecting insect cells, *Spodoptera frugiperda* (Sf9), with three strains of recombinant baculovirus (rBV) with rAAV genome, *rep* gene, and *cap* gene, respectively (Urabe et al., 2002). Later, the system was streamlined to only two rBVs, one carrying the *rep* and *cap* genes, and the other carrying the rAAV genome (Smith et al., 2009). A safety feature of this system was that *rep* and *cap* genes were separated with opposite transcriptional orientations, and the native viral p5 promoter was replaced by heterologous promoters (Allen et al., 1997; Penaud-Budloo et al., 2018; Wright, 2014), so that rcAAV could not be formed. Another system, Monobac, used a single baculovirus encoding AAV *rep/cap* genes and rAAV genome to produce rAAV8, thus reducing baculovirus DNA contamination in the product (Galibert et al., 2021).

Others have sought to create cell lines, which can be rescued to produce rAAV with or without additional virus infection. By integrating the rAAV vector genome and *rep/cap*

genes into HeLa cells, rAAV was produced by infection with Ad (Clark et al., 1995). To eliminate the need for infection of Ad, a HeLa cell line was constructed by integration into its genome an inducible *E1A-E1B* along with rAAV genome, *rep/cap*, *E2*, *E4*, and *VA RNA* genes (Qiao, Li, et al., 2002). The cell line was shown to produce rAAV in the early passages but lost productivity gradually over time. The OneBac system used a *rep-cap* expressing Sf9 cell line, which upon infection with an rBV carrying the rAAV genome, produces the rAAV (Mietzsch et al., 2013). A number of improvements were reported to enhance the quality and productivity of the Sf9 system (Mietzsch et al., 2017; Moreno et al., 2022); however, the possible effect of different posttranslational modifications of rAAV produced in Sf9 cells and in human cells needs to be further investigated (Rumachik et al., 2020). A HEK293 cell-based rAAV production platform with inducible *rep*, *cap*, and helper genes was reported to produce rAAV, but the construction of the cell line was not described (Mainwaring et al., 2021).

We have taken a synthetic biology approach to construct HEK293 cell lines which harbor the rAAV genome and genetic modules that can be induced to express necessary replication, helper, and structural components to produce rAAV (Lee et al., 2022). The approach opened a new paradigm of viral vector production through the creation of stable cell line. However, in those first-generation cell lines, the productivity was low. Here, we report enhancing rAAV productivity in synthetic cell lines through data-driven redesign of genetic modules and tuning of viral gene expression. The method can be readily applied to rAAV of different serotypes with different GOI. Such a stable cell line system will greatly facilitate large-scale manufacturing of rAAV.

3.3 Materials and Methods

3.3.1 Vectors

The rAAV2-GFP genome from pAAV-CAG-GFP (Addgene) was cloned into a Leap-In transposon backbone (ATUM) with a *lacI* repressor gene linked to a puromycin resistance gene driven by a phosphoglycerate kinase (PGK) promoter, and CAG promoter was replaced with a RSV-LacO promoter (Agilent), yielding the inducible rAAV genome module (GM, Figure 3-1A). The construction of the replication module (RM) and packaging module (PM) was previously described (Lee et al., 2022). Modification of RM was made by replacing the GeneSwitch promoter with a TetON promoter and removing GeneSwitch transactivator gene, yielding the new RM (Figure 3-1A). pAAV-CAG-GFP (Addgene), pAAV-RC2 (Cell Biolabs) and pHelper vectors (Cell Biolabs) were used for triple transfection.

3.3.2 Generation of rAAV Producer Cell Lines

HEK293 cells (Cell Biolabs) were cultured in Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 10% fetal bovine Serum (Gibco) and 1X antibiotic-antimycotic (Gibco). GX stable cell lines were generated by cotransfection of three transposon modules with Leap-In Transposase[®] mRNA (ATUM) followed by selection, single-cell cloning, and productivity screening. The workflow is in reference to the previous report and is fully described in Figure 3-2C-E (Lee et al., 2022; Lee et al., 2023).

3.3.3 Cell Culture and rAAV Vector Production

For rAAV production using GX cell lines, cells were seeded at 3.75×10^5 cells per well in 6-well plates or 7.5×10^4 cells per well in 24-well plates. After 24 h, media were replaced with induction media containing 10 $\mu\text{g/mL}$ of doxycycline and 90 $\mu\text{g/mL}$ of

cumate (Sigma Aldrich) unless noted otherwise. For proteasome inhibitor experiments, 1 μ M of MG132 (Invivogen) was added to GX cells at 24 h postinduction (hpi), and cells were incubated with inhibitors for 24 h before media replacement unless noted otherwise. Cells were harvested at 72 hpi. Extraction methods and quantitative polymerase chain reaction (qPCR) assays of intracellular DNA and RNA were previously described (Lee et al., 2022).

3.3.4 rAAV Preparation and Characterization

Preparation of benzonase-treated rAAV crude lysate and titration of encapsidation vector genome (VG) by qPCR, capsids by enzyme-linked immunosorbent assay (ELISA), and transducing units by RM4 assay cells were previously described (Lee et al., 2022). Total intracellular virus genome (TG) copies, including free virus genomes and encapsidated VGs, were measured by qPCR using total intracellular DNA. The full particle content was obtained by dividing the VG by the capsid titer (Grimm et al., 1999).

3.3.5 Targeted Quantitative Proteomics

Sample processing, parameters used for data acquisition by mass spectrometry with parallel reaction monitoring (PRM) method, and data analysis were described previously (Lee et al., 2022). Four new heavy isotope-labeled peptides of VP1, GFP, Rep, and beta-actin proteins were synthesized (Biosynth), and a full list of peptide standards is shown in Table 3-1.

3.3.6 Transmission Electron Microscopy and Cryo-EM

rAAV crude lysate was produced in multiple 150 mm dishes, then purified using POROS CaptureSelect AAVX (Thermo Fisher Scientific) affinity chromatography. rAAV-containing cell lysate was loaded, and the column was subsequently washed with 1 $\times$ TD

buffer (5.3 mM KCl, 137 mM NaCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, 1 mM MgCl₂) (Mietzsch et al., 2020). The adsorbed rAAV was eluted with 0.1 M glycine-HCl (pH 2.6). The rAAV eluate was immediately neutralized with 1 M Tris-HCl (pH 10). For negative-staining TEM, 5 µL of the purified AAV sample was applied to a glow-discharged 400 mesh ultra-thin carbon grid (Electron Microscopy Sciences) and stained by 2% uranyl formate solution (w/v), then blotted dry. For cryo-EM, 3 µL of sample was applied to a glow-discharged R2/1 200 mesh Quantifoil TEM grid (Ted Pella). The blotted dry grid was plunged into liquid ethane, then transferred at the liquid nitrogen temperature to an FEI Tecnai 300kV field emission gun TEM (Thermo Fisher Scientific). The images were recorded on a Gatan 4k slow scan CCD camera or a Gatan Summit K2 direct electron detector camera (Gatan).

Table 3-1. Synthetic heavy isotope-labeled peptides used in PRM.

AQUA Peptide Sequence	Targeted Protein	Coding Transcript	Protein Molecular Function
GIEPTLPNWFVAVT[K_C13N15]	Rep68	Rep68	DNA binding, endonuclease, and helicase activity for genome replication
NVSLPVAHSDA[R_C13N15]	DBP	E4orf6-DBP	Improving processivity of DNA replication
HPPPQILI[K_C13N15]	VP1, VP2, VP3	VP123	Assembly of Capsid at a ratio of 1:1:10
APTEWVIP[R_C13N15]	AAP	VP123	Assembly-activating protein catalyzes capsid assembly process
QLDSGDNPYL[K_C13N15]	VP1	VP123	Possess a phospholipase A2 domain required for AAV infectivity
FSVSGEGEGDATYG[K_C13N15]	GFP	GFP	A fluorescence protein serves as readout
TAPDYLVGQQPVEDISSN[R_C13N15]	Rep	Rep68 or Rep52	(For Rep52/40) Genome encapsidation
AGFAGDDAP[R_C13N15]	Beta-actin		Housekeeping protein

3.3.7 Western Blot

Clarified cell lysate was quantified by BCA protein assay (ThermoFisher Scientific). 10 µg of protein was loaded to each lane of NuPAGE Bis-Tris gels (Thermo Fisher Scientific), blotted membranes were incubated with mouse anti-AAV2 VP1/2/3 (PROGEN) and anti-beta-actin (Cell Signaling Technology) monoclonal antibodies (mAbs), and then incubated with anti-mouse IgG conjugated with alkaline phosphatase mAb (Sigma-Aldrich). Membranes were incubated with Immun-Star™ AP substrate and enhancer (Bio-Rad) for 5 min before chemiluminescence detection.

3.3.8 Flow Cytometry Assay

Cells were fixed with 4% paraformaldehyde, permeabilized with 0.2% Triton X-100, incubated with mouse anti-AAV2 VP1/2/3 (PROGEN) and mouse anti-AAV2 intact particle (PROGEN) mAbs, then incubated with rat anti-mouse IgG1 mAb conjugated with BV421 fluorescence (BioLegend) and rat anti-mouse IgG3 mAb conjugated with AF647 fluorescence (Abcam). Then cells were analyzed on a BD LSR II flow cytometer for GFP, mcherry, BV421, and AF647 fluorescence.

3.4 Results

3.4.1 Controlled Gene Expression of GOI Increased Productivity

We had previously established Pf3 cell line, which synthesized rAAV expressing green fluorescence protein (GFP) as GOI upon induction (Lee et al., 2022). In Pf3 cells, the expression of GFP was driven by a strong constitutive CAG promoter. Its *cap* was encoded by a single intron-less transcript with altered translation initiation, which has been reported to give a more favorable ratio of VP1:VP2:VP3 and denoted as VP123 (Urabe et al., 2002). We used transcriptomic and proteomic analysis to probe potential limiting

factors whose relaxation could lead to enhanced productivity. Pf3 had a transcript level of GFP several times of GAPDH even under uninduced conditions (Figure 3-2A). GFP transcript level increased almost 20-fold upon induction. After viral genome amplification, GFP transcript accounted for 93% of total viral transcripts or 17% of total cellular mRNAs. In comparison, VP123 transcript constituted only 1.64% of viral RNAs or 0.3% of total mRNAs. At the protein level, GFP was also much higher than other viral proteins (Figure 3-2B). Based on the results, we hypothesized that such excessive and wasteful expression of GOI diverted resources, which could have been used for transcription and/or translation of viral genome and proteins and contributed to the low virus productivity. We postulated that minimizing GOI expression in producing cells would increase rAAV production.

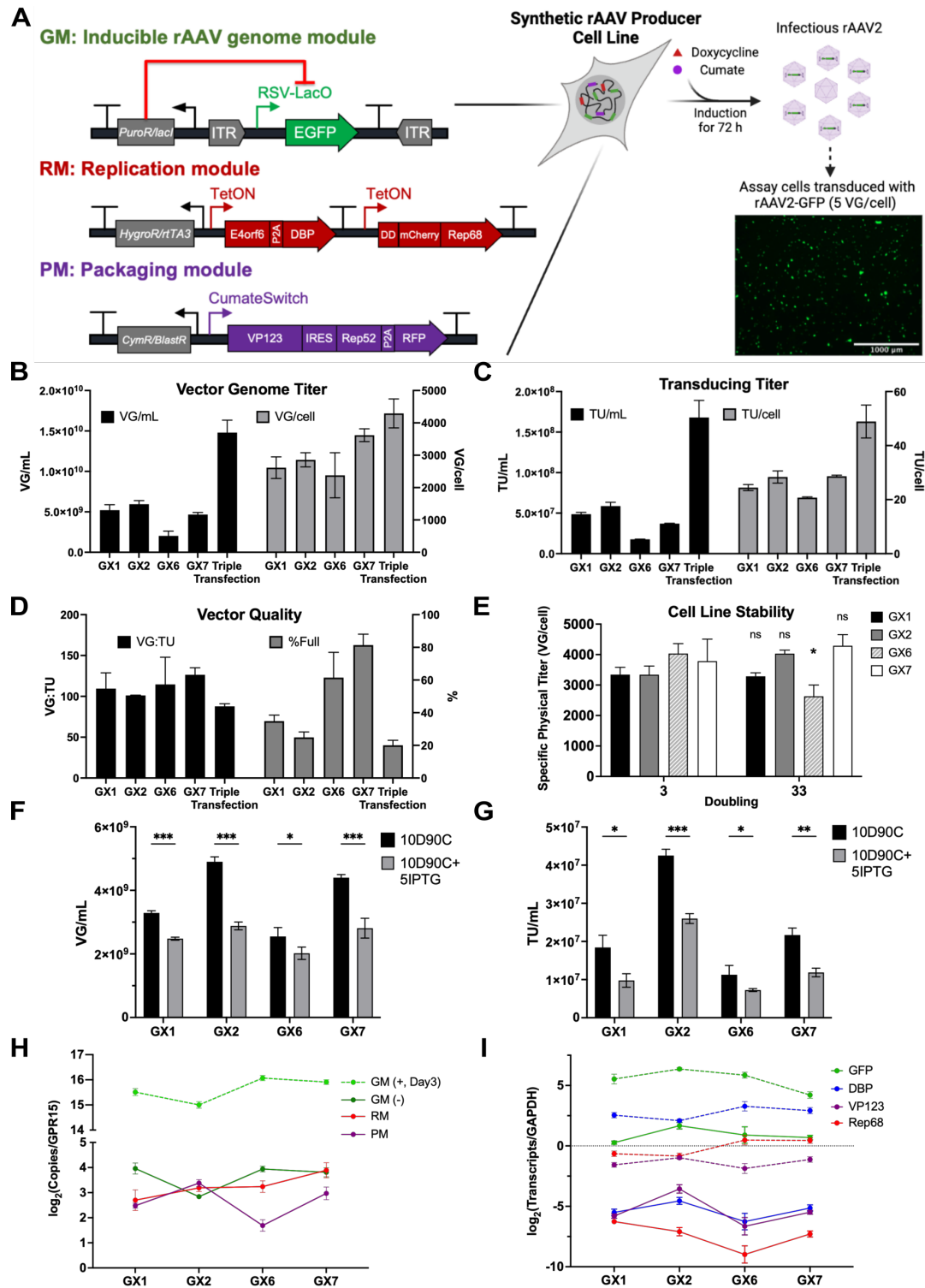


Figure 3-1. Construction and characterization of GX rAAV2 producer cell lines.(A) Gene modules. Symbols: CymR: the Cym Repressor; DD: the destabilization domain; LacO: the lactose (lac) operator; RSV: the Rous sarcoma virus promoter; rTA3: the reverse Tet transactivator gene. (B) Encapsidated vector genome (VG) titers. (C) Transducing unit (TU) titers. (D) The ratio of VG to TU and VG to capsid as vector quality parameters. (E)

Specific rAAV titer in VG/cell at 3 and 33 population doublings. (F, G) VG and TU titers with or without induction of GOI expression by addition of 5 mM of IPTG. (H) Intracellular copy numbers of GM, RM and PM before induction (solid lines) and GM after induction with Dox and cumate for 72 h (+, dash line). (I) Transcript level of viral genes relative to GAPDH before induction (solid lines) and after Dox and cumate induction for 72 h (dash line). rAAV production using GX cell lines was carried out by inducing GX cell lines with 10 μ g/mL of Dox and 90 μ g/mL of cumate (10D90C) for 72 h. For B-D, triple plasmid transfection of HEK293 cells with pAAV-CAG-GFP, pAAV2-RC, and pHelper was included for comparison with GX cell lines. Encapsidated VG titers were measured by qPCR, TU titers by image analysis using RM4 assay cells, and capsid titers by ELISA. For B-G, data represent mean and standard deviation of three independent replicates. For H and I, data represent mean and standard deviation of triplicate qPCR wells. ns: $p > 0.05$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ as determined by an unpaired, two-tailed, two-sample t-test. Multiple-comparisons t-test was assessed by using a false discovery rate (FDR) of 0.05.

We replaced the CAG promoter for GFP in the Genome module (GM) with the Rous sarcoma virus (RSV) promoter controlled by lac operators and repressors (Figure 3-1A) (Wyborski & Short, 1991). The IPTG-inducible lac repressor gene, *lacI*, was placed upstream of the AAV ITR and linked to a puromycin resistance gene via a 2A self-cleaving peptide. Under uninduced conditions, *lacI* is expressed to suppress GFP expression. The replicated and encapsidated genome does not have the *lacI* gene. Upon transducing into a *lacI*-free recipient cell, GFP can be expressed. Additionally, we made a modification to the replication module (RM). The TetON inducible promoter was used to control both Rep68 and helper genes to reduce the size of the RM, eliminating the need of the transactivator for a second inducible promoter. HEK293 cells were transfected with the inducible GM, RM, packaging module (PM), and Leap-In Transposase[®] mRNA. Cells were selected with puromycin, hygromycin, and blasticidin for 2 weeks followed by single-cell cloning. Over 60 clones were screened for high rAAV titer using RM4 assay cells upon induction with 10 μ g/mL of doxycycline (Dox) and 90 μ g/mL of cumate (10D90C) (Figure 3-2C-E) (Lee

et al., 2023). The top four rAAV-producing clones (GX1, GX2, GX6, and GX7) were isolated, expanded, and further characterized.

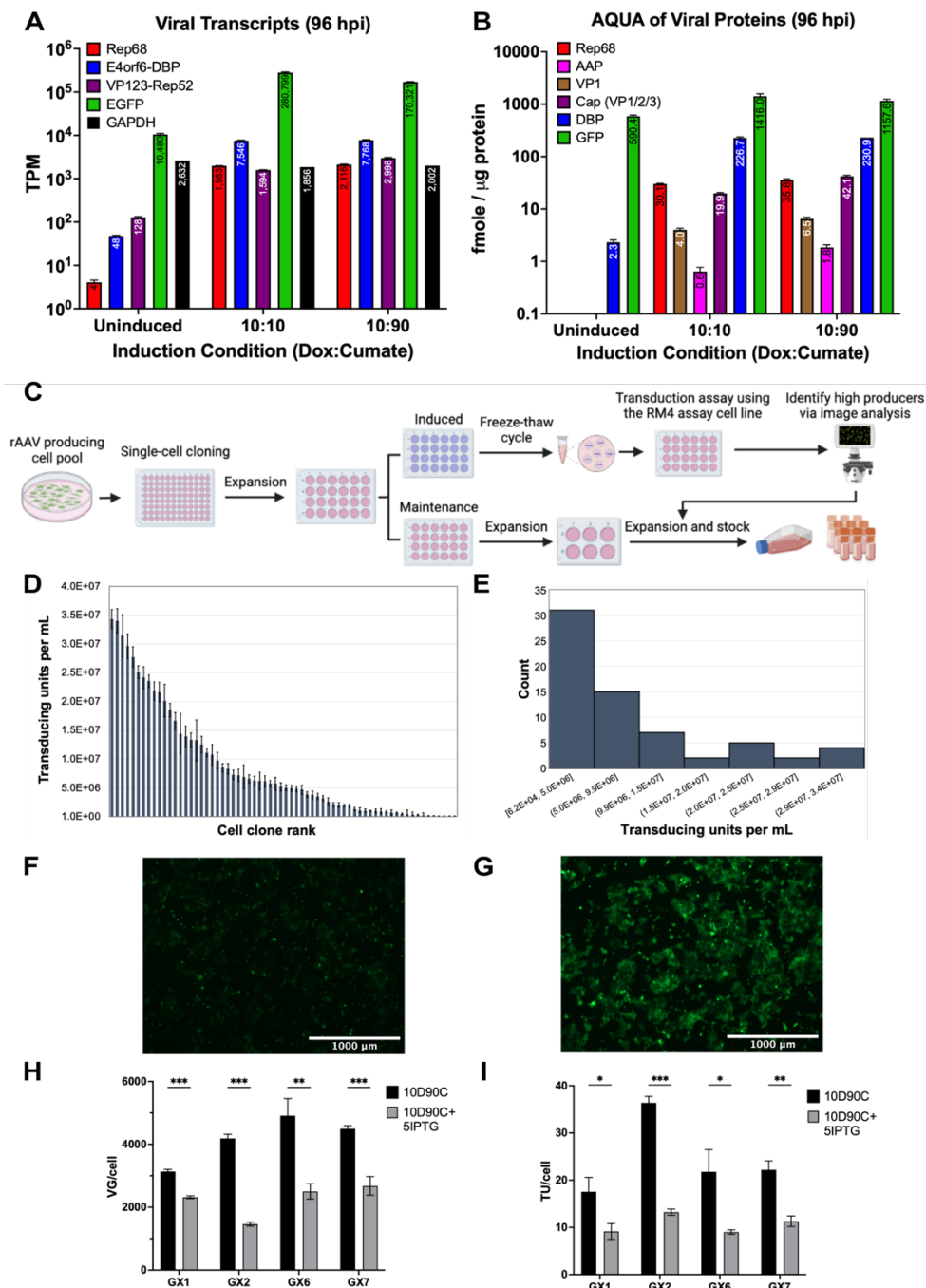


Figure 3-2. Pf3 omics data and cell line development workflow. (A) Transcripts per million of viral genes and GAPDH expressed in Pf3 cells at 96 h post-induction (hpi) measured by RNA sequencing. (B) Absolute quantification of viral proteins expressed in Pf3 cells at 96 hpi measured by mass spectrometry with parallel reaction monitoring (PRM) method. (C) The scheme of rAAV producer cell line development workflow. HEK293 cells were

transfected with three gene modules (i.e., GM, RM, PM, molar ratio = 1:1:1) and Leap-In Transposase[®] mRNA (ATUM) at a 1:1 DNA:RNA weight ratio. After three days of outgrowth, cells were passaged into media containing 2 µg/mL of puromycin (Invivogen), 200 µg/mL of hygromycin B (MilliporeSigma), and 10 µg/mL of blasticidin (Invivogen). After two weeks of selection, single-cell cloning was performed in 96-wells by limiting dilution with a seeding density of 0.5 cells per well. Once reaching confluence, clones were then transferred to 24-well plates for induction and maintenance. After induction for three days, cells were lysed via freeze-thaw cycles, and crude lysate with rAAV was applied to the RM4 assay cells for productivity screening. High producers were expanded for stock. (D) 66 clones were isolated and ranked based on their rAAV titers. (E) Histogram of rAAV titers obtained through induction of each cell clone, and the top 4 clones were used for further analysis. (F) Image of GX2 cells upon induction with 10D90C at 24 hpi. (G) Image of GX2 cells upon induction with 10D90C and 5 mM of IPTG at 24 hpi. (H) VG and (I) TU per cell produced in GX cell lines with or without induction of GOI expression by addition of 5 mM of IPTG. For A and B, Dox was added at 10 µg/mL, cumate was added at 10 or 90 µg/mL, and the mifepristone concentration was kept at 2.5 nM. For C-I, Dox and cumate were added at 10 µg/mL and 90 µg/mL respectively. Data represent mean and standard deviation of three independent replicates. For D, data represent mean and standard deviation of GFP-positive cell counts from multiple images. For H and I, data represent mean and standard deviation of three independent replicates.

GX1, 2, and 7 produced around 5×10^9 encapsidated vector genomes (VG) per mL, and all four clones produced roughly 3000 VG per cell (Figure 3-1B) upon induction by Dox and cumate. A parallel triple transfection using the traditional pAAV-GFP, pAAV-RC2, and pHelper vectors yielded a somewhat higher titer. The transducing titer (assayed using RM4 assay cells and called transducing units [TU]) (Lee et al., 2023) largely followed the trend of vector genome titer (i.e., VG titer) for all four clones (Figure 3-1C). The VG:TU ratio was around 100:1, and the full particle content estimated by VG to capsid ratio was from approximately 25% to 80% (Figure 3-1D). In general, the quality of viral vectors produced using the four clones, in terms of VG:TU and full particle content, was similar to or better than triple transfection. Three out of four clones kept the same productivity after 30 doublings (Figure 3-1E). We evaluated the effect of inducing GOI expression on the rAAV productivity. The *rep*, helper and *cap* genes were induced with or without IPTG induction of GFP expression. With 5 mM of IPTG, producer cells became

much greener, and both VG and TU titers markedly decreased in all four clones (Figure 3-1F-G, Figure 3-2F-I). The result indicated that a high level of GOI expression impeded rAAV production.

The number of copies of each module integrated into the host cell genome, that of amplified viral genome after induction for 72 h, and the transcript level of viral genes were determined via qPCR. The three modules were integrated into the genome in somewhat different numbers for each clone (Figure 3-1H). All four clones had four to eight times more GM integrated than the previously established Pf3 cell line (Lee et al., 2022). Induction at 10D90C caused rAAV genome to replicate to more than 10^5 copies per cell. Although the number of cell clones was small, there appears to be a correlation between copy number of integrated GM to total number of amplified genomes. The genes encoded in RM and PM modules had low levels of leaky expression, but GFP in GM was expressed slightly higher than GAPDH (Figure 3-1I). Upon induction, all viral transcripts became highly expressed. GFP reached a very high level of expression even though its promoter was uninduced because of rAAV genome amplification that could have exhausted the lacI repressor protein.

3.4.2 Dynamics of Viral Gene Expression and Virus Assembly

We characterized the time-course profiles of various viral components after induction. The transcript level of three viral genes, Rep68 and DBP in RM and VP123 in PM, increased rapidly upon induction with Dox and cumate and reached close to GAPDH level by 8 h before leveling off around 16 h in all four cell lines (Figure 3-3A-D). Consistent with the results of Figure 3-1I, DBP was the most highly expressed viral gene in all four cell lines. The transcript levels of Rep68 and the VP123 were several folds lower but still

at high levels similar to GAPDH. GFP transcript followed different dynamics to the viral genes. It started from a high leaky expression level and increased as the viral genome replicated despite the absence of inducer IPTG.

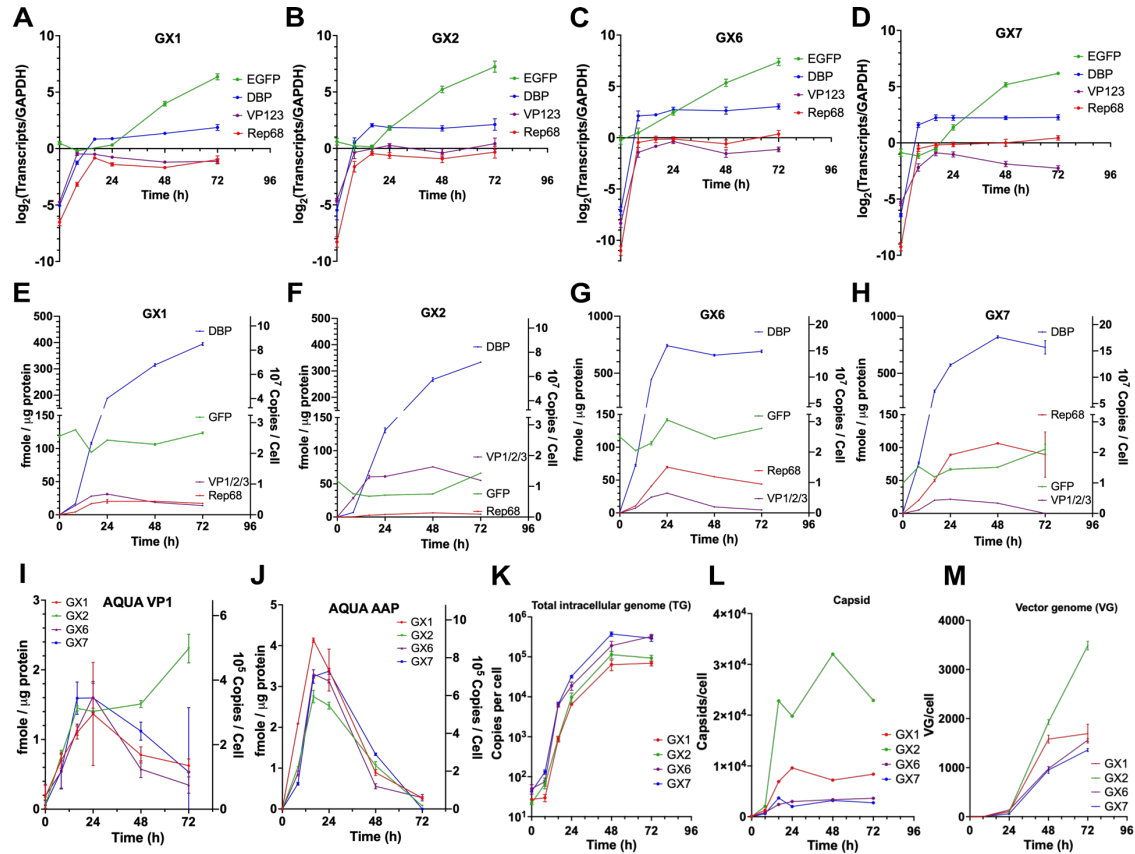


Figure 3-3. Kinetic profiles of viral transcripts and proteins, total intracellular virus genome, assembled capsids, and vector genome in GX cell lines upon induction.(A-D) Transcript level of viral genes relative to GAPDH. (E-J) Absolute quantification (AQUA) of viral proteins measured by mass spectrometry with the parallel reaction monitoring (PRM) method. Protein copies per cell were calculated based on the assumption that total protein from one HEK293 cell is 360 pg. (K) Total intracellular virus genome copies per cell. (L) Assembled capsids per cell. (M) Encapsidated VG per cell. Samples were prepared by inducing GX cell lines with 10D90C. For A-D, K and M, data represent mean and standard deviation of triplicate qPCR wells. For E-J, data represent mean and standard deviation calculated from light to heavy peak area ratios of top three fragment ions in ion chromatograms multiplied by the concentration of peptide standards. For L, data represent mean of duplicate ELISA wells.

Targeted quantitative proteomics was performed on heavy isotope-labeled peptide spiked samples to quantify key proteins. By abundance level, they fell into two broad

groups: Rep68, DBP, VP1/2/3 (quantified using a peptide common to VP1, VP2 and VP3) and GFP in the order of 100 fmol/μg protein (Figure 3-3E-H), while VP1 (quantified using a peptide unique to VP1) and AAP in approximately the level of 5 fmol/μg protein (Figure 3-3I-J). The rise of viral proteins lagged somewhat from the corresponding transcript; the initial fast rise transited to slower increase around 24 h. The relative abundance level and dynamics of different proteins followed similar trends in different clones. The most abundant protein was DBP which had more than 10^7 copies per cell. Notably, VP1 and AAP proteins decreased sharply after peaking around 24 h, along with a similar but less drastic decrease of VP1/2/3. The decrease of VP proteins was also shown in western blot analysis (Figure 3-4). DBP was the only detectable viral protein in uninduced conditions, which had around 5×10^4 copies per cell. A housekeeping protein, beta-actin, had 10^7 copies per cell and remained steady across different conditions (data not shown).

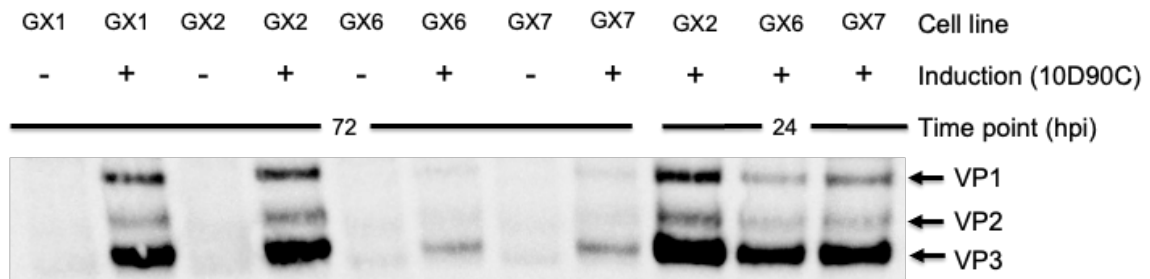


Figure 3-4. Western blot analysis of VP1/2/3 proteins was used to evaluate protein expression of GX cell lines at 24 and 72 h post-induction. Compared to 24 hpi, VP protein expression at 72 hpi was lower, especially in GX6 and GX7 cell lines. The ratio of VP subunits, especially VP1, might affect infectivity of rAAV produced using GX cell lines, so western blot analysis was also used to examine the VP ratio.

Total intracellular virus genomes, including those encapsidated and free ones, reached approximately 10^4 copies at 24 h and increased to around 10^5 copies by 48 h (Figure 3-3K). From 10^6 - 10^7 copies per cell of VP1/2/3 protein, about 10^3 - 10^4 of capsids

were assembled at 24 h (Figure 3-3L). However, the vast majority (>90%) of genome copies and most encapsidated genomes (VGs) were formed after 24 h (Figure 3-3M). Each capsid constitutes 60 VP subunits, hence approximately 3% to 10% of VP1/2/3 proteins were incorporated into capsids (Figure 3-5). The data also suggested that capsid assembly initiated early while VP1/2/3 level was still some way from its peak level around 24 h. Compared with the dynamics of other viral components, genome packaging into capsid was a slow process (Figure 3-3M). Genome encapsidation started early but the rate was slow until total genome copies reached about 10^4 - 10^5 copies per cell. Only a very small portion of total genomes (0.3% - 3%) were encapsidated (Figure 3-5). The data suggested that genome packaging is a potential limiting factor.

Cell line	Total genome (TG/cell)	VG/cell	Capsids/cell	VP1/2/3 protein (copies/cell)	DBP protein (copies/cell)	Rep68 protein (copies/cell)	% TG encapsidated into VG	%VP1/2/3 protein assembled into capsid
GX1	6.90E+04	1693	8380	6.73E+06	8.52E+07	3.70E+06	2.5%	7.5%
GX2	9.33E+04	3485	22900	1.32E+07	7.20E+07	9.21E+05	3.7%	10.4%
GX6	3.23E+05	1561	3630	6.50E+06	1.49E+08	9.46E+06	0.5%	3.3%
GX7	2.94E+05	1359	2750	4.65E+06	1.57E+08	1.93E+07	0.5%	3.5%

Figure 3-5. The specific viral component data of GX cell lines. VP1/2/3 proteins were quantified using a shared peptide representing the sum of all three Cap proteins. VP1, Rep68, and DBP were each quantified by a unique peptide. No unique peptide was available for VP2 and VP3 subunits. TG, VG, and capsids were assayed as described in Materials and Methods. Targeted proteomic data in Figure 2 showed VP1:VP1/2/3 was around 1:20 in GX cells, in the same range as anticipated value of 1:(1+1+10). The theoretical maximum number of capsids that could be assembled was calculated by dividing the peak total VP protein copies by 60, assuming 60 VP1/2/3 units formed a capsid. The percentage of VP1/2/3 actually assembled into capsids was calculated by dividing the actual ELISA-measured capsids by the theoretical maximum number of capsids. The proportion of TG packaged into capsids was calculated by dividing VG by TG. All data presented were obtained from 72 h, except for the VP1/2/3 protein data, which were based on their peak values at 24 h. Only a small fraction of the viral genome replicated were encapsidated, whereas a relatively larger but still small portion of the viral Cap protein VP1/2/3 was assembled into capsids. Higher abundance of VP subunit proteins was correlated to more assembled capsids and higher encapsidated vector genome. Higher levels of Rep and DBP were moderately correlated to higher genome replication and higher full particle contents.

3.4.3 Cell Line-Dependent Optimal Induction Profile for High Productivity

The induction condition used in the preceding experiments was for maximum expression of target genes, which might not be most favorable for rAAV production. A premise of our synthetic biology approach was that the expression profile of viral genes could be modulated by different induction profiles to enhance productivity and product quality. Multiplex screening of the concentration and time profile of Dox and cumate was set up for rAAV production using all four cell lines (Figure 3-6A-D). For GX1 and GX2, the initial induction condition (10D90C) gave a good titer, while for GX6 and GX7, a better titer was obtained by reducing the Dox level, removing Dox at 8 h post-induction (hpi), or delaying addition of Dox. The results show the benefit of a tunable induction system on enhancing the rAAV productivity.

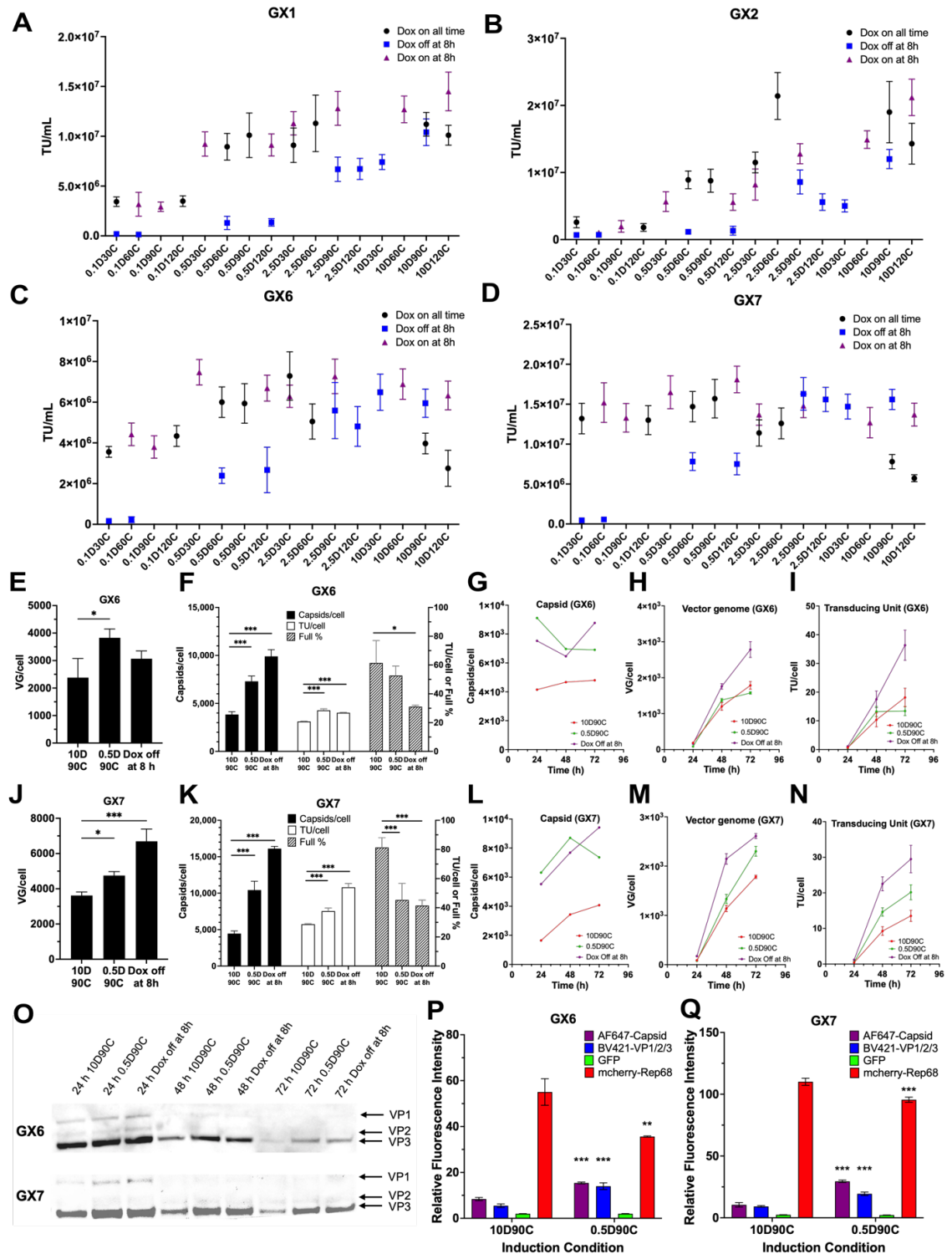


Figure 3-6. Screening of induction conditions on the transducing unit titer.(A-D) Four Dox concentrations (0.1, 0.5, 2.5, 10 $\mu\text{g/mL}$) and three different cumate (30, 60, 90 $\mu\text{g/mL}$) and in conjunction with three different time profiles of Dox exposure, were tested. Inducers were added at 0 h. Dox was present throughout the 72-h period, removed at 8 h or added

at 8 h by replacing the medium. At 72 h, the virus particle was extracted by freeze-thaw method and the transduction unit was assayed using RM4 cells. For GX1 and GX2, the previous induction conditions of 10D90C gave a good titer. While for GX6 and GX7, reducing the level of Dox, removing Dox at 8h, or delaying the addition of Dox gave higher titers than 10D90C. (E-I) and (J-N) show specific data of GX6 and GX7 respectively. Data of three induction conditions, 72 h induction with 10D90C, 0.5D90C, and removal of Dox at 8 hpi are shown. (O) Western blot analysis of VP1/2/3 proteins was used to evaluate protein expression of GX6 and GX7 under the three induction conditions. Both lower dose of Dox and removing Dox at 8 hpi gave higher levels of VP proteins. (P) GX6 and (Q) GX7 induced with 10D90C or 0.5D90C were stained with an antibody against assembled capsid particle (A20) and an antibody against non-assembled VP subunits (B1) to relatively quantify abundance of assembled capsid and non-assembled capsid proteins in cells by flow cytometry. Relative intensity was calculated by dividing mean fluorescence intensity (MFI) by MFI of uninduced GX cell lines stained with antibodies mentioned. Assembled capsids and non-assembled VP subunits abundance were much higher under 0.5D90C condition than under 10D90C for both GX6 and GX7. For A-D, I and N, data represent mean and standard deviation of GFP-positive cell counts from multiple images. For E, F, J, and K, data represent mean and standard deviation of three independent replicates. For G and L, data represent mean of duplicate ELSIA wells. For H and M, data represent mean and standard deviation of triplicate qPCR wells. For P and Q, data represent mean and standard deviation of three independent replicates. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ as determined by a one-way analysis of variance (ANOVA) multiple-comparisons test or an unpaired, two-tailed, two-sample t-test. Multiple-comparisons t-test was assessed by using an FDR of 0.05.

The two new induction conditions, reducing Dox from 10 $\mu\text{g/mL}$ to 0.5 $\mu\text{g/mL}$ (0.5D90C) and removing the 10 $\mu\text{g/mL}$ of Dox at 8 hpi (10D90C-8h) were further compared to the original condition of 10D90C in GX6 and GX7. The results confirmed that 0.5D90C and 10D90C-8h gave higher VG, capsid, and TU titers on both a volumetric basis (Figure 3-7A, B, G, H) and per cell basis (Figure 3-6E, F, J, K). Further characterization was performed on the dynamics of viral gene products under these induction conditions. Consistently, induction with 0.5D90C or 10D90C-8h gave higher levels of assembled capsids, VG, and TU than induction with 10D90C (Figure 3-7C, D, E, I, J, K). However, the total intracellular genome was not markedly different among different induction conditions (Figure 3-7F, L). At the transcript level, expression of Rep68

and DBP decreased greatly with a lower level or removal of Dox as expected (Figure 3-7M, N, P, Q). Since the induction concentration of cumate was the same in all three conditions, VP123 transcript level was similar (Figure 3-7O, R). Nevertheless, the levels of assembled capsid were higher with lower level or removal of Dox (Figure 3-7C, I), so was the level of VP1/2/3 proteins in the western blots and immunofluorescence intensity measured by flow cytometry (Figure 3-6O-Q) suggesting that VP expression might be affected by posttranscriptional regulation.

The kinetic profiles of 0.5D90C and 10D90C-8h were largely similar for both GX6 and GX7. Practically, removing Dox would require the replacement of medium and is difficult to implement in manufacturing. In the following experiments 0.5D90C was adopted for GX6 and GX7.

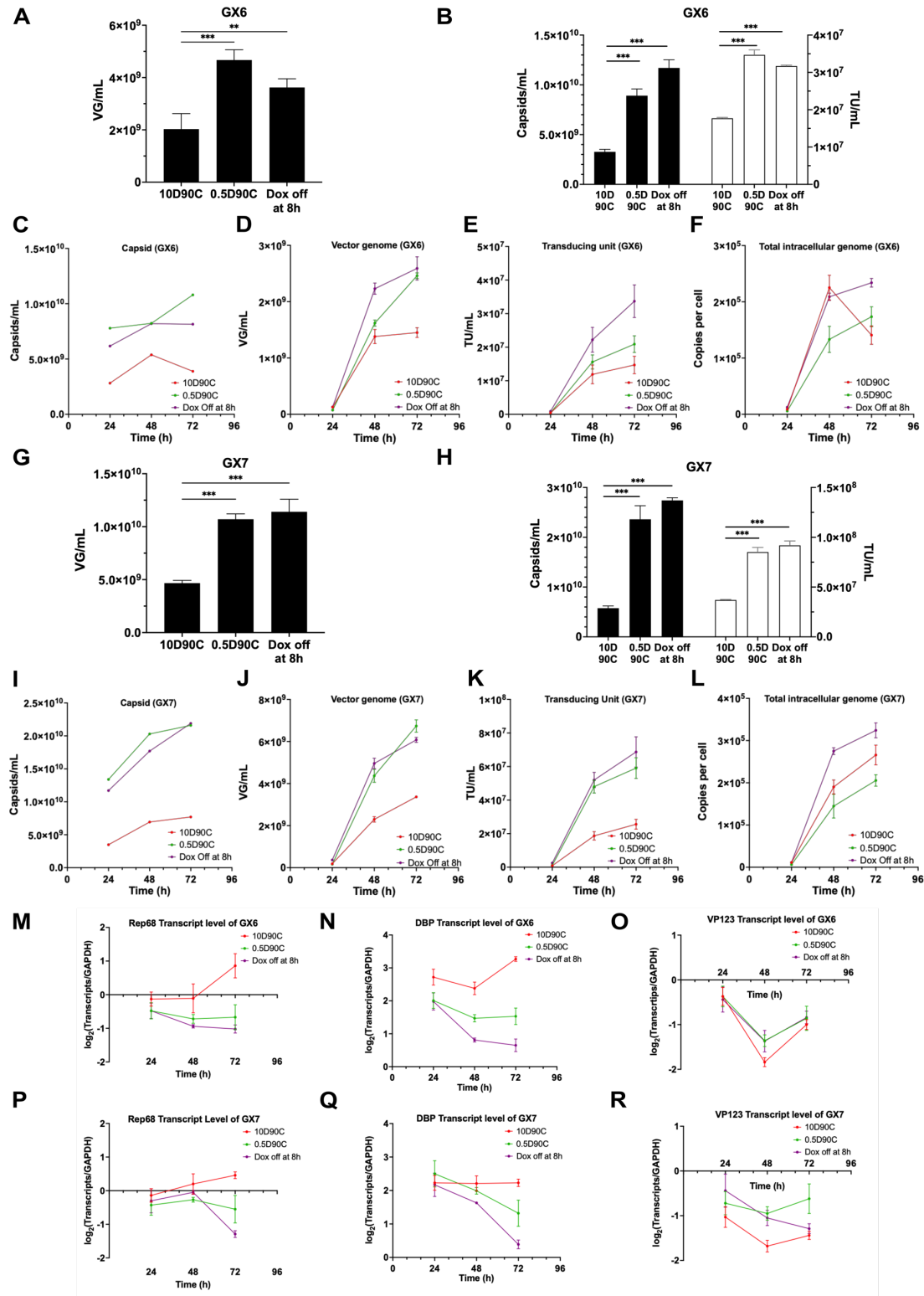


Figure 3-7. Induction conditions affected rAAV production by GX6 and GX7 cell lines. A-F, M-O for GX6, G-L, P-R for GX7 cell line. (A, G) VG, (B, H) capsid and TU titers upon

induction with 10D90C for 72 h, 0.5D90C for 72 h, or 10D90C for 8 h and then with Dox removed for the remainder of 72 h. (C-F, I-L) Kinetic profiles of capsids, total intracellular virus genome, VG and TU under three induction conditions. (M-O, P-R) Relative transcript levels of viral genes. For A, B, G and H, data represent mean and standard deviation of three independent replicates. For C and I, data represent mean of duplicate ELISA wells. For D, E, J, K, and M-R, data represent mean and standard deviation of triplicate qPCR wells. For F and L, data represent mean and standard deviation of GFP-positive cell counts from twelve images. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ as determined by a one-way analysis of variance (ANOVA) multiple-comparisons test.

3.4.4 Inhibition of Proteasomal Degradation Increased Productivity

A striking observation on the dynamic behavior of viral proteins through targeted proteomics was the decline of VP1/2/3 and AAP after reaching peak values around 24 h (Figure 3-3E-J). We postulated that the decrease was caused by proteolytic degradation. In a screening experiment, we tested four proteasome inhibitors on their effect on the infectious rAAV titer under different conditions on the four GX cell lines. An inhibitor was added to cultures at 24 h after induction under optimal induction conditions and incubated for 2 h, 6 h or 24 h before being removed for the remainder of the 72 h. Among the four inhibitors tested, MG132 had a consistent enhancing effect on TU titer using all four cell lines tested (Figure 3-8A-P). Two conditions, MG132 (10 μ M, 6h) and MG132 (1 μ M, 24h) were further evaluated and both showed enhancing effect on the rAAV productivity as compared to the case of without MG132. For GX2 and GX6, both conditions had a similar enhancing effect, but for GX1 and GX7 a higher titer was obtained with MG132 (1 μ M, 24h) (Figure 3-8Q). In subsequent studies, the MG132 at 1 μ M with incubation for 24 h was used for all cell lines.

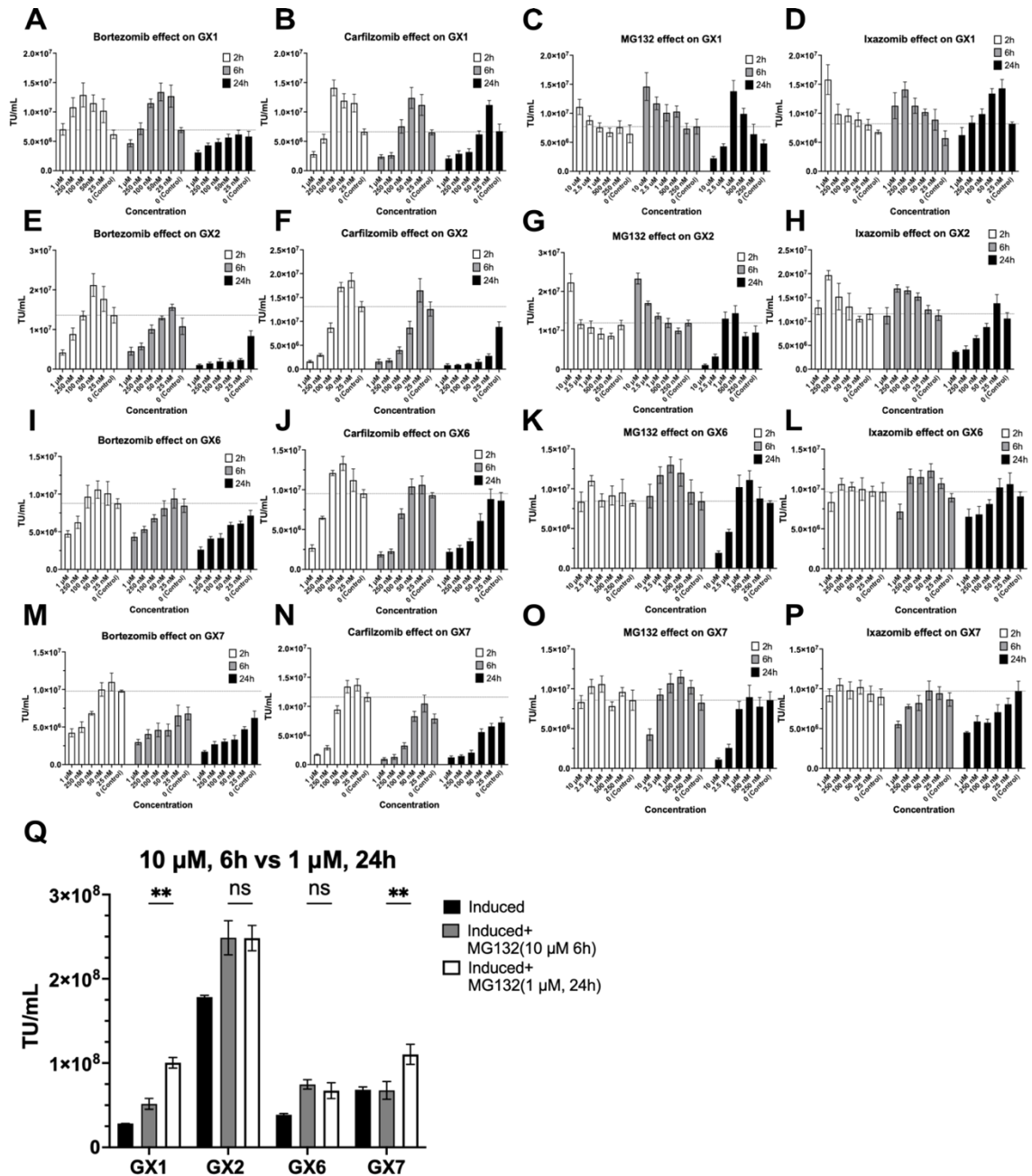


Figure 3-8. Screening of proteasome inhibitors and incubation conditions.(A-P) Cells were first induced under respective optimal induction conditions for 24 h before proteasome inhibitor was added. Different concentrations of Bortezomib (Selleckchem), Carfilzomib (Selleckchem), MG132 (Invivogen), and Ixazomib (Selleckchem) were added to GX cells. After incubation for 2 h, 6 h or 24 h, the media was changed to inhibitor-free and further cultured until 72 hpi. These inhibitor screening experiments were performed with n=1, so more conditions could be tested. The transducing titer was measured using RM4 cells. MG132 had some conditions which enhanced TU titers on all four cell lines. After an effective proteasome inhibitor was identified, the selected inhibitor was tested over a

number of conditions with different concentrations and incubation times in three independent replicates. Shown in (Q) are data with MG132 (10 μ M, 6 h) and MG132 (1 μ M, 24 h). Both conditions had similar enhancing effect on the rAAV productivity. However, for GX1 and GX7, higher titers were obtained with MG132 at 1 μ M incubating for 24 h. For A-P, data represent mean and standard deviation of GFP-positive cell counts from multiple images. For Q, data represent mean and standard deviation of three independent replicates. ns: $p > 0.05$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ as determined by an unpaired, two-tailed, two-sample t-test. Multiple-comparisons t-test was assessed by using an FDR of 0.05.

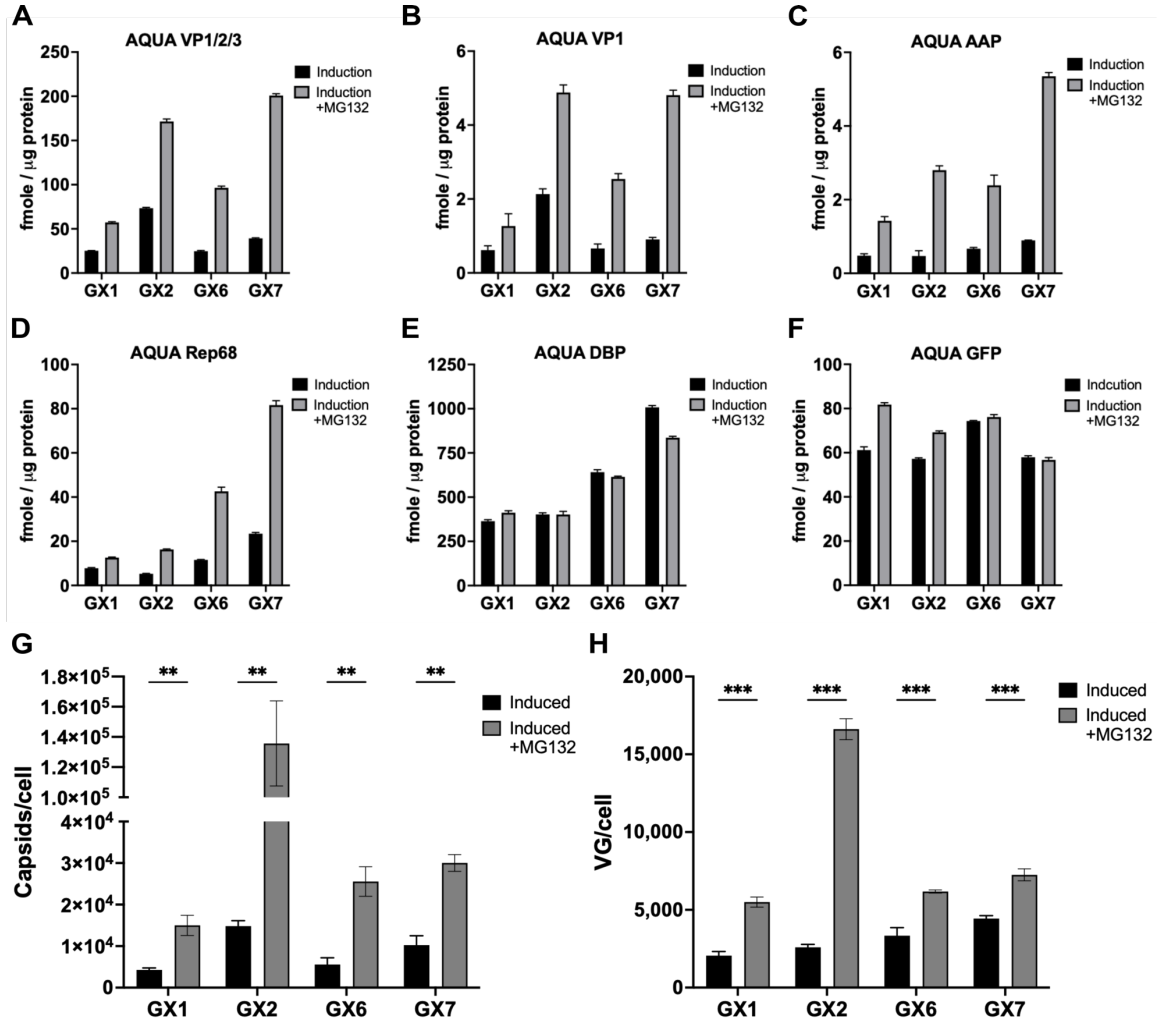


Figure 3-9. Treatment of proteasome inhibitor MG132 increased viral protein abundance and rAAV titers. AQUA of (A) VP1/2/3, (B) VP1, (C) AAP, (D) Rep68, (E) DBP and (F) GFP. (G) Capsids and (H) VG per cell produced in GX cell lines at 72 hpi with or without MG132. GX1 and GX2 were induced with 10D90C, and GX6 and GX7 were induced with 0.5D90C. Proteasome inhibitor, 1 μ M of MG132, was added at 24 hpi for 24 h (MG132, 1 μ M, 24 h) before being removed by medium replenishment with fresh induction media. For A-F, data represent mean and standard deviation calculated from light to heavy peak

area ratios of top three fragment ions in ion chromatograms multiplied by the concentration of peptide standards. For H and I, data represent mean and standard deviation of three independent replicates. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ as determined by an unpaired, two-tailed, two-sample t-test. Multiple-comparisons t-test was assessed by using an FDR of 0.05.

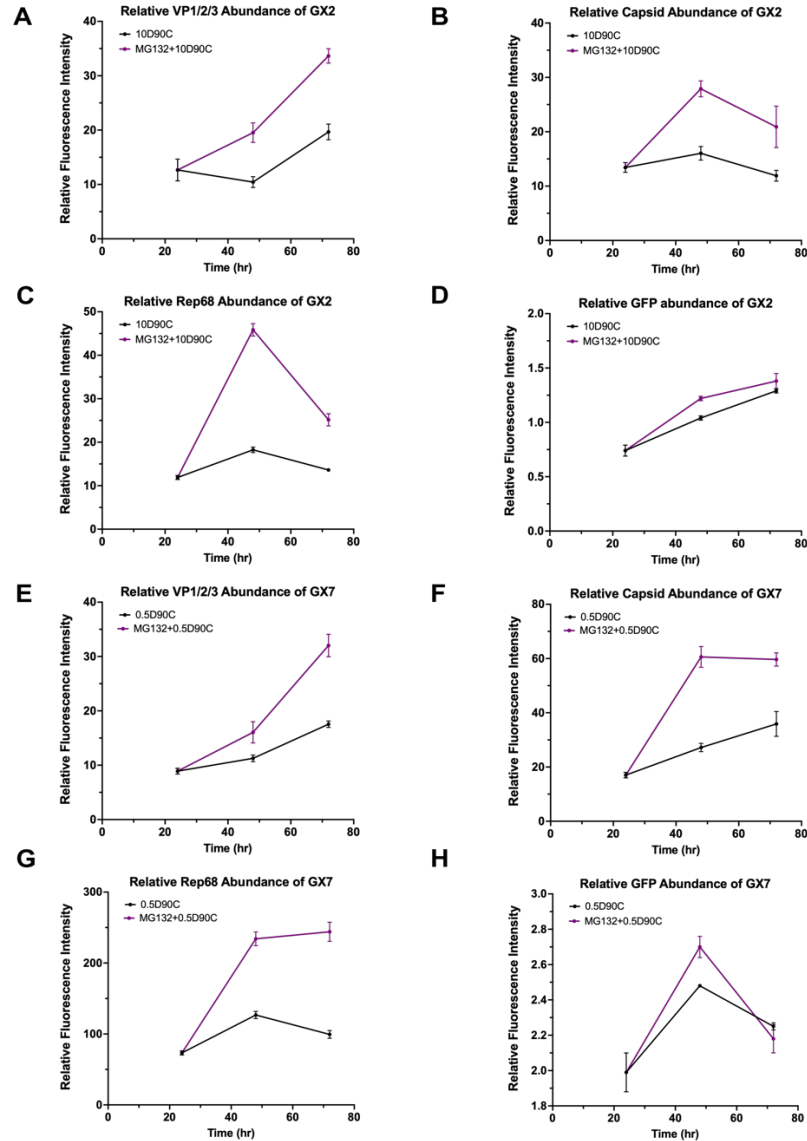


Figure 3-10. The effect of MG132 on viral protein levels was further evaluated by flowcytometry after staining cells with fluorescence-labeled antibodies against VP proteins (A and E) or against assembled capsids (B and F). The addition of MG132 increased the level of both VP1/2/3 proteins and capsids in both GX2 and GX7 under each's respective optimal induction conditions. (C and G) Rep68 proteins were quantified through its linkage to mcherry reporter. (D and H) GFP proteins were quantified also. Relative intensity was calculated by dividing MFI by MFI of uninduced GX cell lines stained with antibodies

mentioned. For A-H, Data represent mean and standard deviation of three independent replicates.

Targeted proteomics quantification on the cell samples subjected to induction and with or without MG132 (1 μ M, 24h) showed significant increase of VP1/2/3, VP1, and AAP proteins with MG132 for all four cell lines (Figure 3-9A-C). Since Rep68 was tagged with a destabilization domain, Rep68 abundance also increased by inhibition of proteasomal degradation (Figure 3-9D). In comparison, the abundance levels of viral protein DBP and viral genome encoded GFP did not increase upon inhibition of proteasome (Figure 3-9E, F). The effect on viral protein abundance was further verified by immunofluorescence intensity measured by flow cytometry (Figure 3-10). With the optimal induction condition and the treatment of MG132, GX2 clones can produce more than 10^4 VG and over 10^5 capsids per cell (Figure 3-9G, H).

3.4.5 Full Particle Content of rAAV Produced

The full particle content produced under optimal induction conditions with or without MG132 was determined by taking the ratio of VG to capsids. While total viral particles increased with MG132 treatment, there was a concurrent increase of empty capsids and hence a marginal decrease in the full particle ratio (Figure 3-11A). Since the full particle content estimated from qPCR and ELISA data had large variations, negative-staining transmission electron microscopy (TEM) and Cryo-EM were used to examine the particle morphology and quantify genome-containing rAAV directly (Subramanian et al., 2019) using samples prepared in two higher producers GX2 and GX7 because of the large quantity of viral particles required for the assay (Figure 3-12). Viral particles produced using GX2 and GX7 were 24% and 56% full, respectively, without MG132. Upon MG132

treatment, it remained at a similar level (26%) for GX2 and dropped moderately (42%) in GX7 (Figure 3-11B).

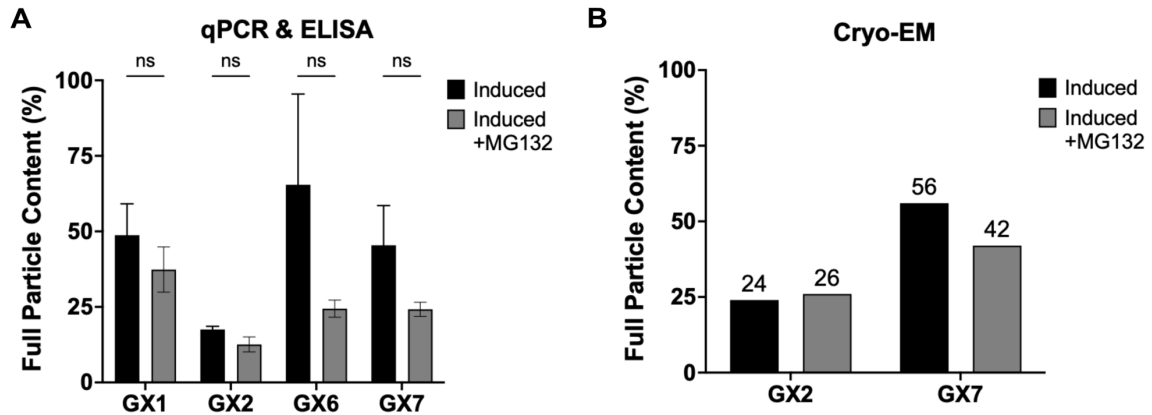
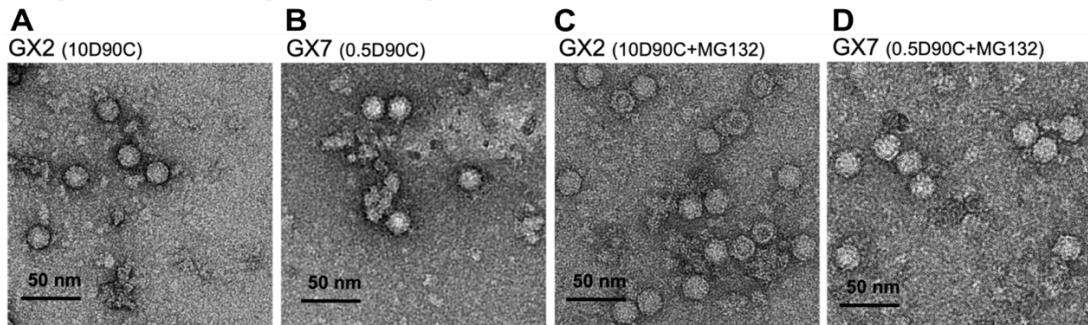


Figure 3-11. Full particle contents produced using GX cell lines under optimal induction conditions with or without MG132 treatment. (A) Full particle content (%) determined by taking the ratio of VG to capsids measured by qPCR and ELISA respectively. (B) Full particle content (%) determined by Cryo-EM analysis. For A, data represent mean and standard deviation of three independent replicates. For B, data represent ratio of full to total particles counted by enumerating particles in multiple Cryo-EM images. ns: $p > 0.05$, as determined by an unpaired, two-tailed, two-sample t-test. Multiple-comparisons t-test was assessed by using an FDR of 0.05.

Negative-staining TEM images



Cryo-EM images

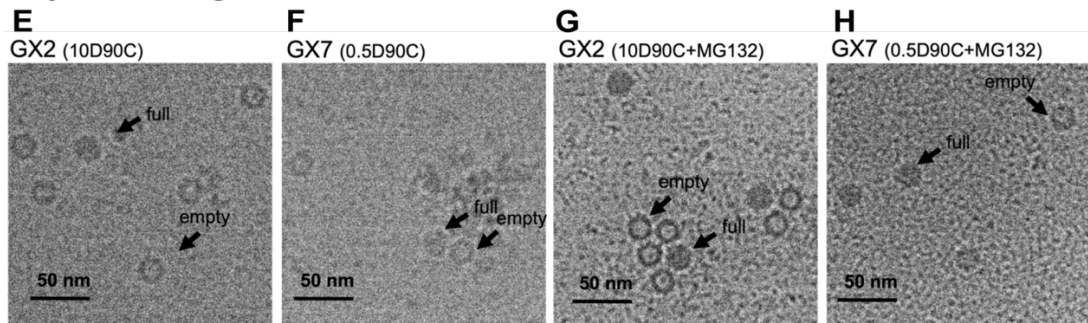


Figure 3-12. Electron microscopic examination and quantification of full particles of the rAAV particles produced using GX2 and GX7 under respective optimal induction conditions with or without MG132. (A-D) Particle dimension and morphology were examined by negative-staining transmission electron microscopy (TEM). (E-H) By Cryo-EM imaging, the difference between full and empty particles could be seen more clearly which gives us more accurate calculations of full particle contents. Thus, full and empty particles produced from high producers, GX2 and GX7, were examined using cryo-EM and enumerated. To generate sufficient quantities of rAAV particles, cells were seeded into multiple 150 mm cell culture dishes at a density of 7×10^6 cells per dish. After one day, GX2 were induced with 10D90C while GX7 were induced with 0.5D90C. For MG132 treatment, MG132 was added to culture at 24 hpi. Virus particles were extracted at 72 hpi using AAVpro Extraction Solution (Takara Bio). The AAV crude lysate was then purified using POROS CaptureSelect AAVX (Thermo Fisher) affinity chromatography. The numbers of empty and full particles in the Cryo-EM images were counted manually.

3.5 Discussion

The design of our synthetic cell lines for rAAV production had two key features: the integration of only minimal essential viral elements into the host cell genome and replacing native viral transcription regulation elements with inducible promoters whose expression can be tuned through exogenous chemical signals. In the current study, we reduced wasteful expression of GOI, isolated new clones with higher copy number of genetic modules, optimized induction profile, and minimized proteasomal degradation of capsid proteins to further advance the productivity. Similar impediment of rAAV production by overexpression of the GOI in transient triple transfection was reported before (Strobel et al., 2015). Using an inducible promoter for GOI, the high-producing GX cell lines isolated indeed had higher numbers of copies of GM integrated into the host genome than the previous Pf cell lines. After induction and genome amplification, the transcript level was still rather high. With our proof-of-concept design of GM, viral genome replication amplified only ITR-flanked regions, but not the *lacI* repressor gene outside the ITR. Hence, after viral genome amplification, the vastly large number of LacO would

quickly exhaust LacI repressor proteins and lead to an increased transcript level of GOI. Further suppression of GOI, for example, by using tightly regulated tissue-specific promoters that have been successfully employed in rAAV targeting different tissues such as muscle (Wang et al., 2008), central neuron system (Tenenbaum et al., 2004), heart and liver (Chen et al., 2013) for GOI expression, may further enhance the productivity.

We compared common features of four GX lines and compared to those of Pf lines to glean possible features that might contribute to higher productivity (Lee et al., 2022). All four GX cell lines had more copies of GM and RM than Pf lines. Upon induction and genome amplification, the difference in total intracellular virus genome between GX and Pf lines widened to more than 10- to 200-fold. This might have contributed to the higher VG titer and vector quality in GX cells. Some common patterns in transcript level of viral genes were seen. DBP transcript and protein levels were very high in all four cell lines. Rep68 and VP123 transcript and protein levels were several-fold lower than DBP, and VP1 and AAP had the lowest levels. However, there was a noticeable difference in Rep68 and VP1/2/3 protein levels among clones. GX2 had somewhat higher VP123 transcript and VP1/2/3 protein levels compared to other cell lines. However, with many variables, including genome replication, capsid assembly, and genome packaging in the production of rAAV, it is hard to assess how much the higher VP protein level contributed to the higher productivity of GX2. Although GX cells had more rAAV genome copies than Pf3, GFP protein was lower as attributed to the use of an inducible promoter for GFP.

Using assay cell line RM4, we surveyed rAAV productivity under different combinations of inducer dose and time profile. While higher levels of Dox induction gave higher rAAV titers in GX1 and GX2, reduced Dox induction increased the productivity in

GX6 and GX7. With 10D, GX6 and GX7 cell lines expressed somewhat higher levels of Rep68 transcript and markedly higher levels of Rep68 protein than GX1 and GX2 (Figure 3-3A-H). The observed enhancing effect on rAAV titer in GX6 and GX7 at reduced expression of Rep68 by using a lower Dox or its removal at 8 hpi may reflect a negative influence of high Rep level on rAAV titer. This was consistent with the report that large Rep proteins inhibited translation and elicited apoptosis (Schmidt et al., 2000; Trempe & Carter, 1988).

Our targeted proteomic study revealed possible degradation of AAP and VP1/2/3 proteins after reaching their peak values in all four GX cell lines (Figure 3-3E-J). AAP facilitates cotransport of VP proteins to the nucleoli and promotes VP protein stability and capsid assembly by oligomerization (Maurer et al., 2018; Sonntag et al., 2010). Without AAP, VP proteins are subject to degradation through several pathways, including proteasomal degradation in the unfolded protein response (UPR) as part of host cell response to virus infection (Balakrishnan et al., 2013; Prasad & Greber, 2021). Adding a proteasome inhibitor in rAAV transduction has been shown to have an enhancing effect on transgene expression (Jennings et al., 2005; Mitchell & Samulski, 2013). Inhibiting proteasome-mediated degradation during rAAV production led to a higher level of AAV2 capsid proteins and boosted rAAV production. This enhancing effect of rAAV productivity was attainable with a high level of MG132 for a shorter period or by using a low MG132 concentration throughout the induction. The titer-enhancing effect of MG132 provided additional evidence of the role of host cell-initiated degradation in rAAV production.

With enhanced induction conditions and MG132 treatment, the productivity in terms of VG was approaching that of traditional triple transfection. For future

manufacturing, one anticipates that the host cells will be suspension-growth adapted to allow for high-density cultures. The use of a chemical inhibitor in manufacturing may not be desirable. One may consider the possibility of knocking out the responsible protease genes.

The full particle content ranged from 25% to over 50%. The value was comparable to or higher than the 30% reported in the literature (Penaud-Budloo et al., 2018). VG:TU ratio is often used as an estimate of rAAV infectivity. Its value can be highly variable between production methods or even preparation lots (François et al., 2018; Mohiuddin et al., 2005). rAAV produced from GX cells across different batches of production had VG:TU about 100, a value comparable to prior reports (Zeltner et al., 2010). While the number of cell lines and rAAV production kinetic studies is still small, a general trend seems to emerge. Increasing the VP1/2/3 protein level results in higher capsid titers as well as VG titers. However, their increase is not proportionate; instead, the capsid titer increases more than the VG titer, leading to a drop in full particle content. An optimal production condition will need to consider both VG titer and full particle content.

We had previously reported proof-of-concept rAAV-producing cell lines created using a synthetic biology approach by employing only essential gene sets for rAAV synthesis and placing them under inducible promoter control (Lee et al., 2022). Here we report our success in further enhancing productivity through cell engineering and optimization of gene expression dynamics and culture conditions. The productivity of the GX cell lines approaches that of commonly used triple plasmid transient transfection. Also, only a tiny portion of total genome encapsidated, thus one may speculate that the productivity can possibly be enhanced by the higher expression of *cap* genes and increasing

the efficiency of viral genome packaging into capsids. Further exploration of helper viral components such as Ad noncoding VA RNA, which inhibits PKR-mediated antiviral response (Vachon & Conn, 2016) and other Ad E4 proteins (e.g., E4orf3 and E4orf6/7) (Huang & Hearing, 1989; Soriano et al., 2019) which modulate host responses along with tempering of host cell responses to rAAV replication may have much enhancing effect on the productivity. With further enhancement, the use of stable cell lines for rAAV manufacturing may become a reality.

4 Elevating Recombinant Adeno-Associated Virus Titer and Quality by Modulating Viral Gene Expression Dynamics in High-Producing Synthetic Cell Lines

This work is adapted from the manuscript in preparation for submission. Min Lu and Wei-Shou Hu conceived and designed the study and prepared the manuscript. Min Lu conducted experiments, analyzed data, and collaborated with others as follows: Han-Jung Kuo and Yu-Chieh Lin contributed to suspension adaptation, and Wen Cai expanded isolated clones.

4.1 Summary

Recombinant adeno-associated virus (rAAV)-mediated gene therapies have achieved numerous clinical successes and are now widely used as *in vivo* gene delivery vectors. However, the substantial quantities required for broad applications present formidable manufacturing challenges. To address this issue, we have developed cell lines integrated with essential viral genes, which can be induced to produce rAAV without plasmid transfection or virus transduction. Through iterative design-construct-test cycles, we've showcased the potential of this synthetic cell production system. In this study, we have placed a greater emphasis on the testing phase of the development cycle, utilizing systems characterization with transcriptomic and proteomic tools to identify means to enhance productivity and product quality. Low *cap* gene expression and resulting low capsids for genome encapsidation was dealt by sequential transfection with a modified packaging module encoding Cap proteins, elevating the rAAV titer to 10^5 vector genome per cell. High full particle content was achieved without sacrificing total encapsidated particles through tuning of inducer concentrations and timing. Integrated view of viral

element kinetics highlighted the importance of balancing genome amplification and capsid formation on full particle content. Cell lines were then adapted to scalable serum-free suspension culture. Our synthetic and systems biology approach enables iterative enhancements and learning in rAAV production, thereby making the use of an inducible producer cell line for rAAV production an industrially relevant process.

4.2 Introduction

Adeno-associated virus (AAV) is a small capsid virus with a 4.7 kb-long single stranded DNA genome. The small genome contains only two genes, *rep* and *cap*, which generates six transcripts and nine proteins (Maurer & Weitzman, 2020). The two genes are sandwiched in between two inverted terminal repeats (ITR). It is named because its infection of a cell results in latent infection unless it is co-infected with a helper virus, such as Adenovirus. For gene therapy applications, the sequence in between two ITRs are replaced with a therapeutic gene. Because of its small genome size, the payload gene of interest that can be carried by the recombinant AAV (rAAV) is also limited. However, rAAV vector is safe as it cannot replicate by itself unless a helper virus is present. It exhibits tissue tropism across its various serotypes, is episomally persistent for long-term expression after transduction into host cells, and is suitable for *in vivo* applications. It hence is an attractive gene delivery system and has become one of the most widely used *in vivo* gene delivery vehicles (Mendell et al., 2021). rAAV has contributed to a number of clinical successes and gained numerous regulatory approvals, including Luxturna for retinal dystrophy, Zolgensma for spinal muscular atrophy (Au et al., 2022), Hemgenix for hemophilia B, Elevidys for Duchenne muscular dystrophy, and Roctavian for hemophilia A.

rAAV is typically applied in very large doses. For *in vivo* systemic administration through intravenous injection, around 10^{13} - 10^{14} physical particles per kg of patient's weight (e.g., Zolgensma: 1.1×10^{14} , Hemgenix: 2×10^{13} , Elevidys: 1.3×10^{14} , Roctavian: 6×10^{13} VG per kg) (Au et al., 2022; Paulk, 2020) are needed. Large doses of rAAV required place a huge challenge on manufacturing. The dose may be reduced through targeted administration to specific organs or tissues, such as down to 1.5×10^{11} VG for the eye. However, in general applications, a very high dose is the norm. To meet the demand, large-scale production of up to 100 liters is often needed to generate enough rAAV for one patient (Merten, 2016); thus, innovation and enhancement of rAAV manufacturing are urgently needed.

Since rAAV genome lacks AAV genes, to produce rAAV, the *rep* and *cap* genes must be supplied in *trans*, with helper genes also provided. A simple and most commonly used system is thus transient transfection of HEK293 cells with multiple plasmid vectors carrying rAAV genome, *rep/cap* genes, and helper viral genes (i.e., *E2A*, *E4*, *VA RNA*) (Xiao et al., 1998). HEK293 cells additionally carries E1A and E1B helper genes in their genome. Recent developments in dual-plasmid systems have contributed to higher transfection efficiency and productivity (Tang et al., 2020; van Lieshout et al., 2023). The method is straight forward to implement, but it needs additional production process(es) of multiple plasmid generation (Shupe et al., 2022).

Others have employed viral vectors, such as the herpes simplex virus-BHK cell system or the baculovirus-insect cell system, as carriers of the rAAV genome and AAV *rep-cap* genes, with helper functions provided by the carrier viral vectors (Clément et al., 2009; Joshi et al., 2021). A HeLa packaging cell line which harbors rAAV genome and *rep-cap* genes has been generated (Chadeuf et al., 2000). Upon transduction of Adenovirus,

rAAV is produced. These systems have relatively better scalability; however, they all require additional process(es) for virus production, the establishment of these virus stocks, and the implementation of measures to safeguard against the potential presence of viruses and helper viral nucleic acids and proteins (Dobrowsky et al., 2021).

These rAAV production systems can produce up to 10^5 rAAV particles per cell on an order of magnitude basis (Merten, 2016; Ou et al., 2023). However, the viral vector produced generally has a large percentage (50-90%) of empty particles without the genome and payload gene (Penaud-Budloo et al., 2018). These empty capsids might serve as decoy, contributing to neutralization of the host humoral immune response (Mingozzi & High, 2013). However, a high abundance of empty particles has been reported to cause poor liver transduction due to binding competition (Gao et al., 2014). Furthermore, even payload-less empty particles can enter target cells, leading to antigen presentation recognized by capsid-specific T cells (Mingozzi & High, 2013) and potentially elicits immune response (Sun et al., 2003). Therefore, increasing concerns were being raised on the high empty particle content in rAAV for clinical applications. In laboratory scales, empty particles are commonly separated from full particles by ultracentrifugation exploiting their slightly different buoyant densities, but this method is not scalable (E. Ayuso et al., 2010). Anion exchange chromatography has been successfully employed for large-scale separation of empty particles from full particles, relying on the difference in isoelectric point between full and empty particles (Dickerson et al., 2021; Qu et al., 2015). An ideal new system for rAAV production should be robust, free of virus infection or plasmid transfection, with a high productivity and high full particle content.

We have hypothesized that while a high level of synthesis of VP proteins (which assemble into capsids) and rAAV genomes (which are packaged into capsids to form full particles) is important for a high productivity, the kinetics of capsid formation, genome replication and packaging are also critical in determining the full-to-empty particle ratio. To allow for stoichiometric and kinetic manipulation of key components of rAAV, we have taken a synthetic and systems biology approach to construct producer cell lines with the following features: (1) instead of using the native AAV *rep* gene, we used the Rep68 and Rep52 coding sequences while removing native promoters and placed them under the inducible TetOn and CumateSwitch control, respectively; (2) a destabilization domain is linked to Rep68 to minimize its possible cytotoxicity caused by leaky expression while unstabilized; (3) we placed *cap* gene under CumateSwitch control (removing p40) while also removing the intron and altering the translation initiation codon to adjust the VP1:VP2:VP3 ratio to be in-line with the natural ratio in a capsid (Urabe et al., 2002); (4) two helper genes under the inducible TetOn control, E4orf6 and DBP, were included, relying on HEK293 cells to provide E1A and E1B functions; (5) these genetic elements were grouped into three modules based on their roles (genome, replication, and packaging) and placed under TetOn and CumateSwitch control to allow for possible independent induction of genome replication and capsid protein expression; (6) a LacSwitch inducible promoter controls the expression of the payload gene, which inevitably is present in a very high copy number, to minimize its competition for cellular resources. To facilitate systems analysis, we quantified viral protein content by proteomic assays, evaluated transcriptome and viral transcripts by RNAseq, and established an assay cell line to allow for speedy measurement of infectious rAAV produced (Lee et al., 2023). With the synthetic biology

design and the systems analysis tools we generated cell lines that produced rAAV at specific titers near the level of traditional triple plasmid transfection in three rounds of design-build-systems optimization/analysis cycle (Lee et al., 2022; Lu et al., 2023).

Herein, we report another round of design-build-systems optimization/analysis cycle to further boost the productivity to be on a par with or higher than traditional methods. We further exploited our system's versatility in dynamic control of induction to tune the productivity and full particle content. The presented inducible rAAV producer cell line is scalable, robust, and cost-effective without the use of plasmid transfection and virus infection. High-titer rAAV with a high full particle content produced using the inducible cell line demonstrates its readiness for large-scale gene therapy manufacturing.

4.3 Materials and Methods

4.3.1 Vector Construction

The construction of the genome module (GM), replication module (RM), and packaging module (PM) was previously described (Lee et al., 2022; Lu et al., 2023). Modification of the PM for sequential transfection involved replacing the CymR repressor gene with the coding sequence of the blue fluorescence protein (BFP), yielding the new PM-BFP.

4.3.2 Generation of rAAV Producer Cell Lines

Parental GX cell lines were previously developed by co-transfection of three transposon vectors (i.e., GM, RM, PM) with Leap-In Transposase[®] mRNA 1 (ATUM) followed by selection, single-cell cloning, and productivity screening (Lu et al., 2023). GX2 and GX6 cells were seeded at 7.8×10^4 cells per cm². Sequential transfection of GX2 and GX6 cell lines with PM-BFP and Leap-In Transposase[®] mRNA 2 (ATUM) was

conducted. At 48 hours post-transfection, cells were transferred to T75-flasks for expansion. Cells were passaged for 1 week to dilute out excess donor plasmid. Subsequently, cells were sorted on a BD FACSAria II Cell Sorter with a 100 μm nozzle. BFP-positive cells were sorted into 15-mL tubes. The productivity of two cell pools, GX2-SX and GX6-SX, were evaluated by transduction assay. Afterwards, GX6-SX pool was selected for further single-cell cloning performed in 96-wells by limiting dilution.

4.3.3 Cell Culture and rAAV Production

For rAAV production using synthetic producer cells, cells were seeded at 3.9×10^4 cells per cm^2 . After 24 hours, media were replaced with induction media containing 10 $\mu\text{g/mL}$ of doxycycline and 90 $\mu\text{g/mL}$ of cumate (Sigma Aldrich) unless noted otherwise. Cells were harvested at 72 h post-induction (hpi). Methods for extracting intracellular DNA and RNA, as well as conducting qPCR and qRT-PCR assays for gene and viral genome copy number, and transcript analysis, respectively, have been previously described (Lee et al., 2022). The copy number of the integrated genetic module was calculated in reference to the copy number of GPR15 in the HEK293 cell genome. The transcript level presented was relative to the transcript level of GAPDH.

4.3.4 rAAV Preparation and Characterization

The preparation of benzonase-treated rAAV crude lysate and the titration of encapsidated vector genomes (VG) and total intracellular virus genomes (TG) by qPCR, the titration of capsids by ELISA, and titration of transducing units (TU) by RM4 assay cells (Lee et al., 2023) were previously described (Lee et al., 2022; Lu et al., 2023). By qPCR analysis on the total intracellular DNA, the copies of TG were calculated in reference to the copy number of GPR15 in the HEK293 cell genome. By qPCR analysis on the

benzonase-resistant VG, the copies of VG were calculated based on the standard curve generated through a series of serial dilutions of the plasmid standard. The full particle content was obtained by dividing the VG by the capsid titer (Grimm et al., 1999).

4.3.5 Targeted Quantitative Proteomics

Sample processing, parameters used for data acquisition by mass spectrometry with parallel reaction monitoring (PRM) method, and data analysis using Skyline were previously detailed (Lee et al., 2022). Light-to-heavy peptide ratios were calculated with peak areas in ion chromatograms, and then, ratios were used to determine endogenous protein concentrations (fmole/ μ g protein). New heavy isotope-labeled peptides for MAAP and E4orf6 quantification were synthesized (Biosynth), and a complete list of peptide standards is provided in Table 4-1. Estimation of the number of protein molecules per cell is detailed in Table 4-2.

Table 4-1. Sequences of heavy isotope-labeled peptides used in the targeted proteomics assay for absolute quantification (AQUA).

AQUA Peptide Sequence	Targeted Protein	Coding Transcript
DFLTEW[R_C13N15]	Rep68	Rep68
NVSLPVAHSDA[R_C13N15]	DBP	E4orf6-P2A-DBP
EELVIL[R_C13N15]	E4orf6	E4orf6-P2A-DBP
HPPPQILI[K_C13N15]	VP1, VP2, VP3	VP123-IRES-Rep52
APTEWVIP[R_C13N15]	AAP	VP123-IRES-Rep52
GLVLPGY[K_C13N15]	VP1	VP123-IRES-Rep52
LLGATSDEQSS[R_C13N15]	MAAP	VP123-IRES-Rep52
FSVSGEGEGDATYG[K_C13N15]	EGFP	EGFP
AGFAGDDAP[R_C13N15]	ACTB	ACTB
AILGGS[K_C13N15]	Rep68, Rep52	Rep68 / VP123-IRES Rep52

Table 4-2. Estimation of the number of protein molecules per cell.

Total protein (pg) per cell	360
--------------------------------	-----

Avogadro constant (Molecules per mole)	6×10^{23}
AQUA data (fmole per μg protein)	1
Conversion formula	$1 \frac{\text{fmole}}{\mu\text{g protein}} \times 10^{-6} \frac{\mu\text{g}}{\text{pg}} \times 10^{-15} \frac{\text{mole}}{\text{fmole}} \times 360 \frac{\text{pg}}{\text{cell}}$ $\times 6 \times 10^{23} \frac{\text{molecules}}{\text{mole}} = 216,000 \frac{\text{molecules}}{\text{cell}}$

4.4 Results

4.4.1 Identification of Limiting Factors in Synthetic rAAV Producer Cell Lines

We reported the development of synthetic rAAV producer cell lines by integration of a minimal sets of essential viral genes for rAAV production. Three genetic modules comprising an essential set of coding sequences (CDS) were integrated into cell genome: Genome module (GM) has an ITR-flanked EGFP driven by a LacSwitch promoter, which represses EGFP expression and contributes to enhanced rAAV productivity (Figure 4-1A). Replication module (RM) contains CDS of helper proteins, E4orf6 and DBP, and the AAV Rep68 protein tagged with mCherry and a destabilization domain (DD), all driven by TetOn promoters. Packaging module (PM) has an intron-less *cap* gene, VP123, linked with the Rep52 CDS via IRES driven by a CumateSwitch promoter. Removing the intron in *cap* gene allows for expression of VP1, VP2, and VP3 subunits at a correct stoichiometry without relying on alternative splicing. Four cell lines, named GX1, GX2, GX6, and GX7, were obtained after selection with puromycin, hygromycin, and blasticidin and single-cell cloning through limiting dilution (Lu et al., 2023).

Although rAAV productivity was enhanced by using identified optimal induction conditions (i.e., 10 $\mu\text{g/mL}$ of doxycycline and 90 $\mu\text{g/mL}$ of cumate (10D90C) for GX1 and

GX2, 0.5 $\mu\text{g/mL}$ of doxycycline and 90 $\mu\text{g/mL}$ of cumate (0.5D90C) for GX6 and GX7), fewer than 10^4 vector genome (VG) per cell were produced using GX cell lines. To further enhance the rAAV titer, viral gene expression in GX cells and triple transfection system was analyzed and compared at both the transcript and protein levels (Figure 4-1B, C). In the qRT-PCR result, compared to the triple transfection system, GX cell lines had similar levels of Rep68 and helper transcripts but a much lower level of the Cap (VP123) transcript (Figure 4-1B). Consistently, protein abundance determined by targeted proteomics revealed that Cap proteins (i.e., VP1/2/3 subunits) and AAP, which are responsible for capsid assembly, in GX cell lines were several folds lower than in the triple transfection system (Figure 4-1C). In addition, we revisited and compared the abundance of total intracellular virus genome (TG) and capsids to VG titers and found out that more than 10^5 copies of TG available for packaging while only 10^3 - 10^4 capsids available that limited VG titers in GX cell lines (Figure 4-1D). Together, supported by viral gene expression and viral component data, low *cap* gene expression stood out as the possible limiting factor for low capsid production and VG titer.

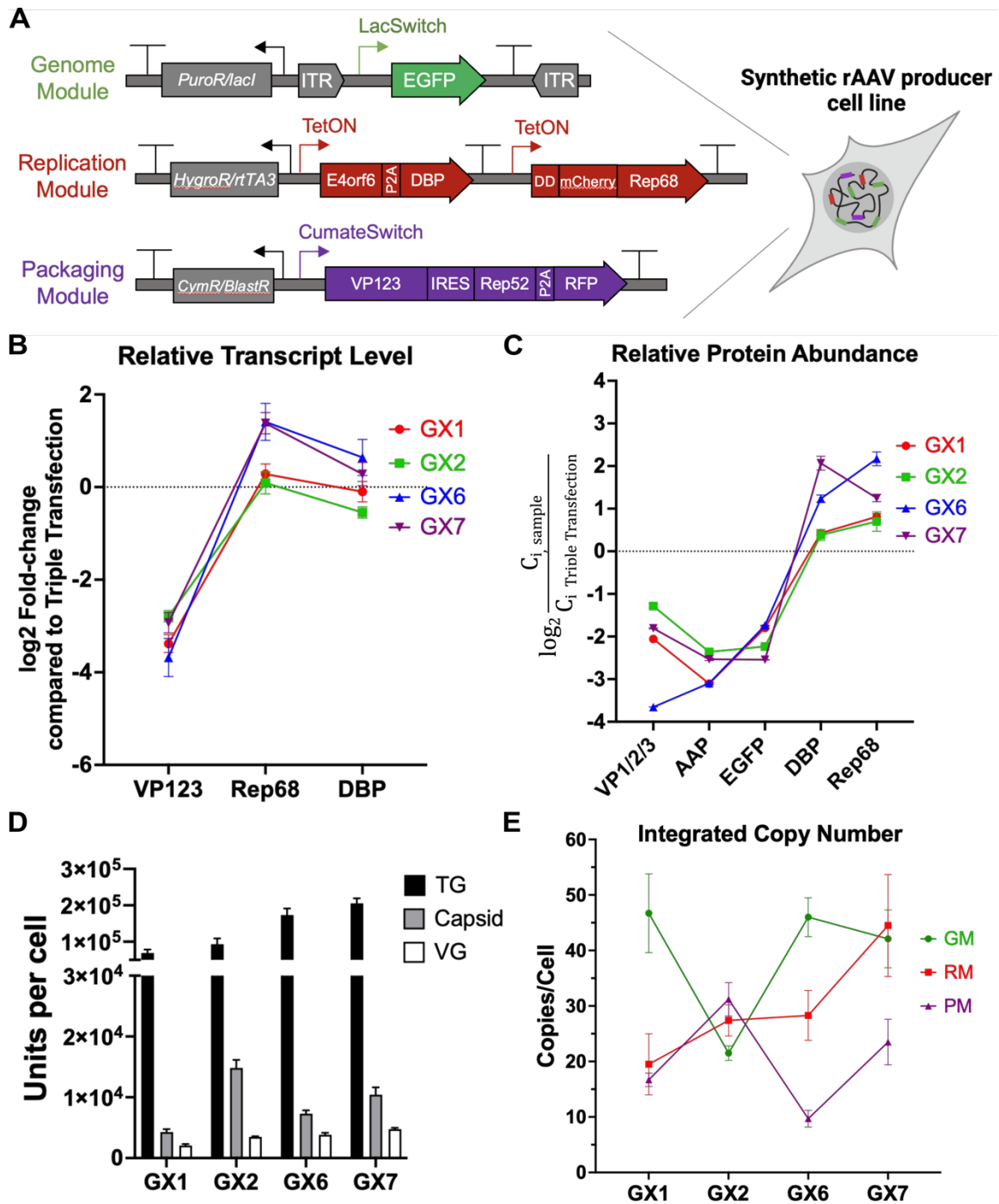


Figure 4-1. Identifying limiting factors in GX rAAV2 producer cell lines. (A) Gene modules integrated in the HEK293 cell genome. Symbols: lacI: the Lac repressor gene; rtTA3: the reverse Tet transactivator gene; CymR: the Cym Repressor; DD: the destabilization domain. (B) Transcript levels of viral genes relative to GAPDH compared to those observed in the triple transfection system. (C) Abundance of viral proteins measured by targeted proteomics with the parallel reaction monitoring (PRM) method compared to those measured in the triple transfection system. (D) Total intracellular virus genomes, VG and capsid titers produced by GX cell lines. (E) Copy numbers of integrated

GM, RM and PM in GX cell lines. Encapsidated VG titers and TG were measured by qPCR and capsid titers by ELISA. For B-D, samples were prepared by inducing GX1 and GX2 with 10 µg/mL of Dox and 90 µg/mL of cumate (10D90C) for 72 h and inducing GX6 and GX7 cells with 0.5 µg/mL of Dox and 90 µg/mL of cumate (0.5D90C) for 72 h. For B, D, E, data represent mean and standard deviation of triplicate qPCR wells or duplicate ELISA wells. For C, data represent mean and standard deviation calculated from light to heavy peak area ratios of top three fragment ions.

4.4.2 Sequential Transfection with The Packaging Module Substantially Increased

Capsid and VG titers

Increasing the copy number of a transgene has been shown to enhance transgene expression in mammalian cell lines. Thus, we introduced more copies of the PM encoding VP1/2/3 proteins and AAP into GX cell lines, aiming to increase capsids and boost rAAV productivity. GX2 and GX6 cell lines were selected for sequential transfection. GX2 had a higher copy number of PM (around 30 copies) and the highest capsid titer (Figure 4-1D, E). GX6 had a lower copy number of PM (around 10 copies) but a higher level of TG (around 2×10^5 copies). A greater enhancement could be achieved by integrating more copies of the PM into GX6, as this would increase the availability of capsids for packaging with abundant virus genomes. A modified packaging module (PM-BFP) with a blue fluorescence protein (BFP) gene as a sorting marker was constructed. GX2 and GX6 cell lines were transfected with PM-BFP and Leap-In Transposase[®] 2 mRNA at a 1:1 weight ratio in a 6-well plate (Figure 4-2A). At 48 h post-transfection, cells were transferred to T75 flasks for expansion. After reaching confluency, BFP-positive cells were sorted by fluorescence-activated cell sorting (FACS) to give two cell pools, GX2-SX and GX6-SX, respectively (Figure 4-2A and Figure 4-3A).

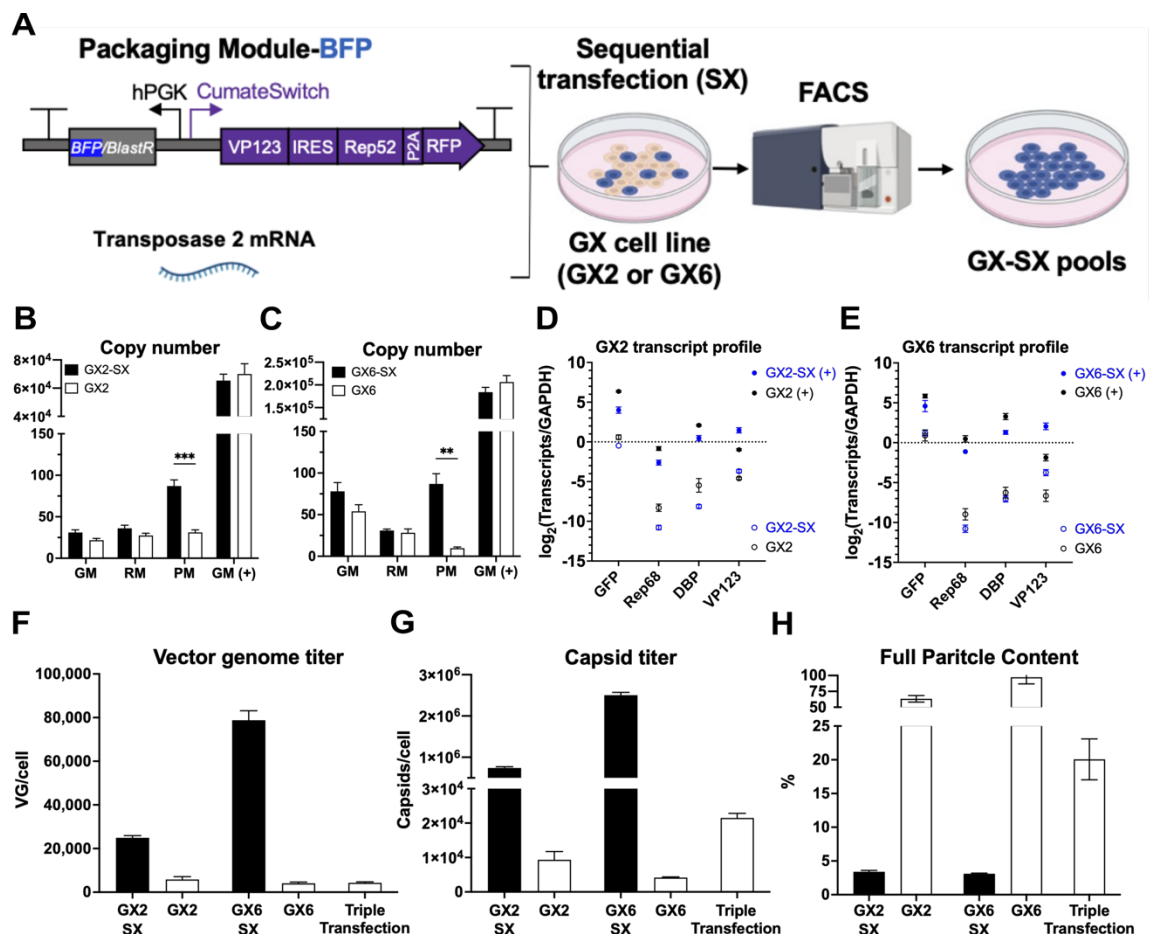


Figure 4-2. Development of rAAV2 high-producing cell pools by sequential transfection. (A) The schematic workflow of sequential transfection of GX2 and GX6 cells with a modified packaging module (PM-BFP) with blue fluorescence protein (BFP), and collection of BFP-positive cells by fluorescence-activated cell sorting (FACS). (B) Copy numbers of integrated GM, RM and PM and intracellular copy number of GM after induction with Dox and cumate for 72 h (+) in GX2 and GX2-SX cells and (C) in GX6 and GX6-SX cells. (D) Transcript levels of viral genes relative to GAPDH before induction (empty circle) and after Dox and cumate induction (+) for 72 h (solid circle) in GX2 and GX2-SX cells and (E) in GX6 and GX6-SX cells. (F) Encapsidated vector genome (VG) titers. (G) Capsid titers. (H) The full particle content calculated by dividing the VG titer to the capsid titer. rAAV production was carried out by inducing GX2 and GX2-SX cells with 10 μ g/mL of Dox and 90 μ g/mL of cumate (10D90C) for 72 h and inducing GX6 and GX6-SX cells with 0.5 μ g/mL of Dox and 90 μ g/mL of cumate (0.5D90C) for 72 h. Encapsidated VG titers were measured by qPCR and capsid titers by ELISA. For B-E, data represent mean and standard deviation of triplicate qPCR wells. For F-H, data represent mean and standard deviation of three independent replicates. ns: $p > 0.05$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ as determined by an unpaired, two-tailed, two-sample t-test. Multiple-comparisons t-test was assessed by using a false discovery rate (FDR) of 0.05.

Sorted cell pools and parental cell lines were induced with 10D90C, previously experimentally determined optimal induction conditions for rAAV production, for GX2-SX and GX2, and 0.5D90C for GX6-SX and GX6. At 72 hours post-induction (hpi), cells were harvested and lysed by freeze-thaw cycles. The released virus in the supernatant was quantified using assay cells, RM4 (Figure 4-3B) (Lee et al., 2023). Both GX2-SX and GX6-SX pools had substantially higher transducing unit (TU) titers than each's parental cell line (Figure 4-3C).

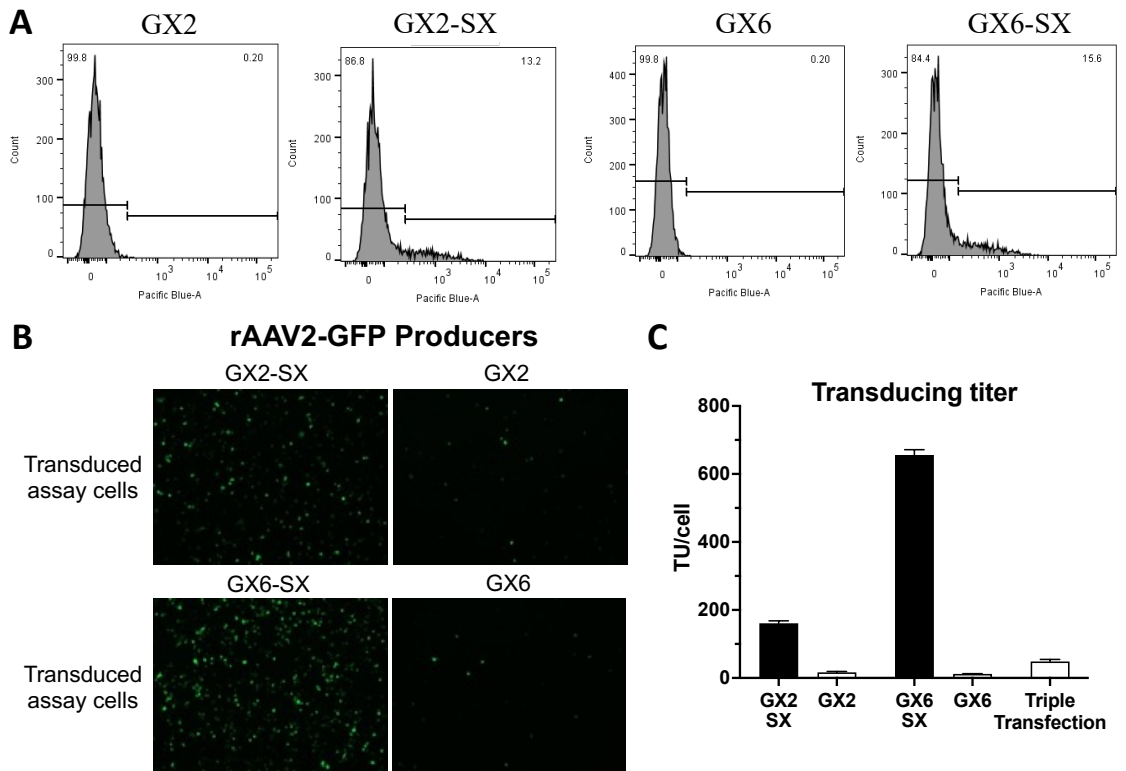


Figure 4-3. Sorting BFP-positive cells by FACS and evaluating rAAV productivity of SX cell pools by the transduction assay.(A) Flow cytometry population histograms of BFP-intensity upon sequential transfection (SX) of GX2 and GX6 cells with PM-BFP. (B) GFP fluorescence images of RM4 assay cells transduced with rAAV2-GFP obtained from parental cell lines, GX2-SX and GX6-SX pools. (C) Transducing titers measured by RM4 assay cells. For C, data represent mean and standard deviation of three independent replicates.

We then measured the number of PM integrated into the genome and the VP123 transcript level by qPCR analysis on intracellular DNA and qRT-PCR analysis on cDNA, respectively. Both GX2-SX and GX6-SX had more integrated copies of PM while similar copies of GM (both before and after induction) and RM as compared to GX2 and GX6 (Figure 4-2B, C). Furthermore, upon induction, both pools also had 5- to 15-fold higher VP123 transcript levels than each's parent (Figure 4-2D, E). Interestingly, transcript levels of Rep68, DBP, and GFP were somewhat lower in SX pools than the parents.

Since subsequent single cell cloning entailed much effort, we further characterized the GX2-SX and GX6-SX pools for each's productivity and an important quality index, i.e., full particle content. Both pools had much higher VG titers than parental cell lines and the triple transfection system (Figure 4-2F). The capsid titers increased more than 100-fold as compared to parental cell lines (Figure 4-2G). More capsids contributed to higher VG titers, but the excessive capsids produced in new pools also resulted in low full particle content (Figure 4-2H).

4.4.3 Isolation and Characterization of High-Producing Cell Clones

Limiting dilution was performed on the GX6-SX pool due to its higher productivity. 51 clones were isolated and screened for rAAV productivity using RM4 assay cells (Figure 4-4A, B). More than two thirds of the clone had less than half of the productivity of top clones. Top four clones, upon induction with 10D90C, all produced around 10^5 VG per cell (Figure 4-5A). The number of copies of PM integrated among the four clones varied somewhat (Figure 4-5B). Capsid titers and full particle content also varied among clones (Figure 4-5C, D). GX6A and GX6D had more integrated copies of PM, potentially resulting in higher capsid titers and lower full particle content compared to GX6B and

GX6C (Figure 4-5B-D). Interestingly, all four clones had slightly or moderately fewer integrated copies of PM than the GX6-SX pool (Figure 4-2C and Figure 4-5B). GX6A cell line were then passaged for about 30 doublings, and no significant decrease in the rAAV titer on a per-cell basis was seen (Figure 4-6A). GX6B cell line was also stable, as illustrated in the suspension cell adaptation (Figure 4-6B, C). Reversal to the adherent state gave similar rAAV productivity.

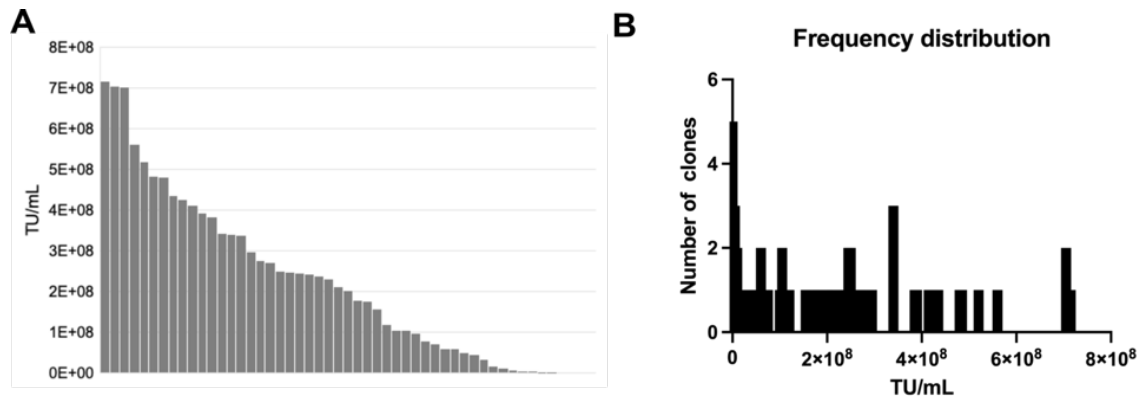
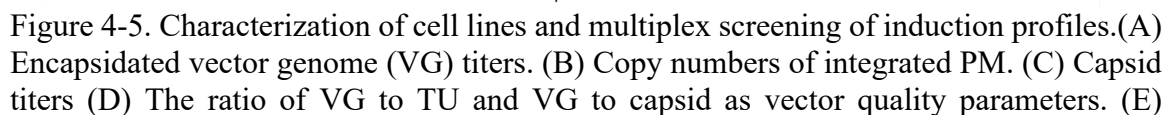


Figure 4-4. Single-cell cloning and productivity screening of clones by transduction assay.(A) 51 clones were isolated and ranked based on their TU titers obtained through induction of each cell clone. (B) Histogram of TU titers, and the top 4 clones were used for further analysis. Data represent mean of GFP-positive cell counts from multiple images.

Two clones, GX6A and GX6B, were selected for further investigation because of their somewhat higher titers. Absolute quantification (AQUA) of viral proteins using targeted proteomics were performed on GX6A, GX6B, and the parent GX6 under the induction condition of 0.5D90C. Proteins encoded in PM (i.e., Cap, VP1, AAP, MAAP, and Rep52) increased highly as compared to the parent GX6, while those encoded in RM and GM remained at similar levels as parent (Figure 4-5E). With more integrated copies of PM, Cap proteins (i.e., VP1/2/3) became the most abundant viral protein, followed by the DBP and Rep52 protein. The VP1 protein also increased proportionally and accounted for roughly 9% of total VP1/2/3 proteins. The two orders of magnitude increase in Cap proteins



Absolute quantification (AQUA) of viral proteins measured by mass spectrometry with the parallel reaction monitoring (PRM) method. (F) Screening of induction conditions on the VG and capsid titers. Two Dox concentrations (0.5, 10 $\mu\text{g/mL}$) and three different cumate (10, 30, 90 $\mu\text{g/mL}$) concentrations, and in conjunction with three different time profiles of cumate addition (0, 6, 24 hpi), were tested. VG and capsid titers were normalized to the respective lowest titer. For A-D, data represent mean and standard deviation of three independent replicates. For E, data represent mean and standard deviation calculated from light to heavy peak area ratios of top three fragment ions in ion chromatograms multiplied by the concentration of peptide standards. For F, * represent induction conditions further tested (i.e., 10D10C, 10D90C, 0.5D90C).

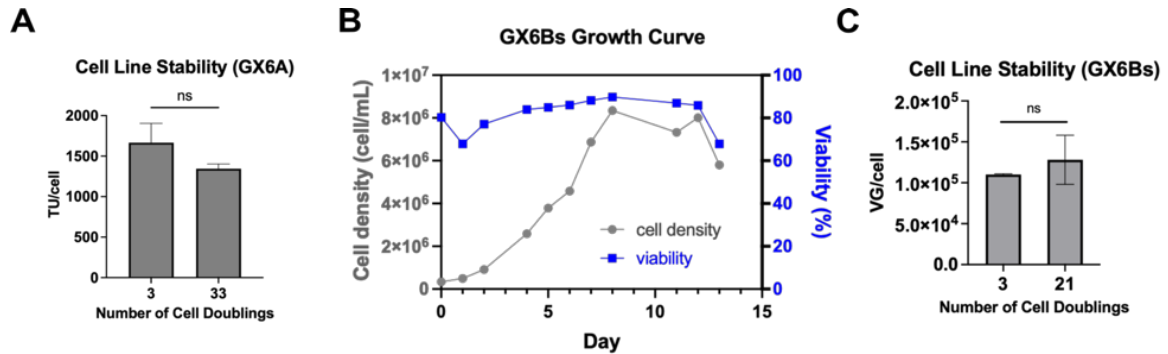


Figure 4-6. Cell line stability and suspension adaptation. (A) Specific rAAV titer in TU/cell at 3 and 33 population doublings. rAAV production was carried out by inducing GX6A cell line with 10D90C. For large-scale culture, suspension growth is preferred over adherent cultivation. GX6B cells were thus adapted to grow in suspension. GX6B were trypsinized and transferred to shaker flasks with a 30 mL working volume in a protein-free Celer-S001 medium at a cell concentration of 5×10^5 cells/mL. After a period of static cell density with little apparent proliferation, cells began to proliferate. On Day 19, GX6B cells were harvested and resuspended into two new shaker flasks and began to expand its population for building up cell stocks (named GX6Bs). (B) GX6Bs grew to 8×10^6 cells/mL in a batch culture with a doubling time of about 32 hours during the exponential growth phase. GX6Bs cells from day 6 of a batch culture were resuspended in DMEM (in which GX6B grew) containing 10% FBS and induced with 10D90C. In a few hours, they readily to adhere to culture plates. VG harvested at 72 hpi was assayed. (C) GX6Bs produced rAAV with a titer on a par with GX6B cells under the same culture and induction condition. GX6Bs, when grown as adherent cells, retained the productivity of GX6B cells.

4.4.4 Multiplex Screening for Improving Vector Quality

A premise of our synthetic biology approach to constructing multiple-inducible cell lines for rAAV production was the versatility of independent tuning of the synthesis of different viral components to optimize productivity and product quality. Boosting the copy number of PM increased the Cap protein level, and the TU and VG titers. However, the

increased capsid titer was not met by a proportional increase in viral genomes packaged into capsids. A large excess of capsids resulted in a high level of empty particles. We tested our premise of controlling the product quality and screened induction conditions using GX6A and GX6B for induction conditions that could produce high-titer rAAV with a favorable full particle content.

We had observed previously that rAAV genome replication had a slower initiation as its rapid amplification did not start until a sufficient quantity of Rep68 proteins and rAAV genome templates have accumulated after 24 hpi (Lu et al., 2023). In contrast, Cap protein expression initiated rapidly from the integrated PM and assembled into capsids relatively early. In our modular design, Rep68 in the RM is under TetOn control and Cap and Rep52 in the PM is under CumateSwitch control. Hence, Tet induction indirectly controls rAAV genome replication, while cumate induction has more effect on capsid production. To tune induction condition, we can manipulate both the inducer concentration and the time profile of inducers. We postulated that an early onset of Dox induction is necessary to have a high amplification of genome after 24 hpi. Hence, we tuned Tet induction by adjusting Dox level. We then dialed cumate induction down by either reducing its concentration or delay the onset of its induction.

A multiplex experiment to survey the effect of a wide-range of induction conditions, including three cumate concentrations (10, 30, 90 $\mu\text{g/mL}$), two Dox concentrations (0.5, 10 $\mu\text{g/mL}$), and three cumate induction initiation times (0, 6, 24 hpi), on rAAV titer and full particle content was conducted (Figure 4-5F). The results are plotted as normalized capsid titers vs. VG titers. Note that the VG/capsids ratio also represent full particle content. Granted that rAAV production is a complex process, and its outcome is affected by

multiple interactive variables non-linearly. Plotting the data in two dimensions is a gross simplification. Nevertheless, some qualitative trends were seen.

For GX6A (red symbols in Figure 4-5F), a general trend of cumate dose could be better seen. Starting from the high capsids/cell side of the plot, reducing cumate concentration or delaying its addition by 24 h shifted capsids/cell to the lower side. This was first occurred without much being affected by Dox level of 0.5 or 10. The lack of a clear relationship between Dox dose and VG also reflects the complex and relatively indirect relationship between Dox, the inducible replication genes, and VG. VG formation required genome replication as well as packaging. For GX6B (blue symbols in Figure 4-5F), VG titers were mostly lower than GX6A. A general trend of cumate could be seen, high concentration data points are more to the right, low concentrations and delayed induction points are to the left, albeit that the data points clustered in a narrower range and data points overlapped substantially. For GX6A, 10D10C posed to be an induction condition that balanced a good full particle content without much sacrifice of total encapsidated particles (Figure 4-5F). For GX6B, 10D90C appeared be a good trade-off induction condition.

4.4.5 Shifting Balance between Genome Replication and Capsid Formation

The effect of altering induction conditions and shifting the balance between genome replication and capsid formation was examined further under the original inductions of 10D90C and 0.5D90C for two cell lines and a reduced cumate level of 10D10C. To shed light on the kinetic behavior of various key elements, the study was performed over different hours post-induction (hpi). Data obtained at 72 hpi are shown in Figure 4-7. The complete time-course profiles of the viral elements under three induction concentrations for GX6A and GX6B are shown in Figure 4-8 and Figure 4-9 respectively.

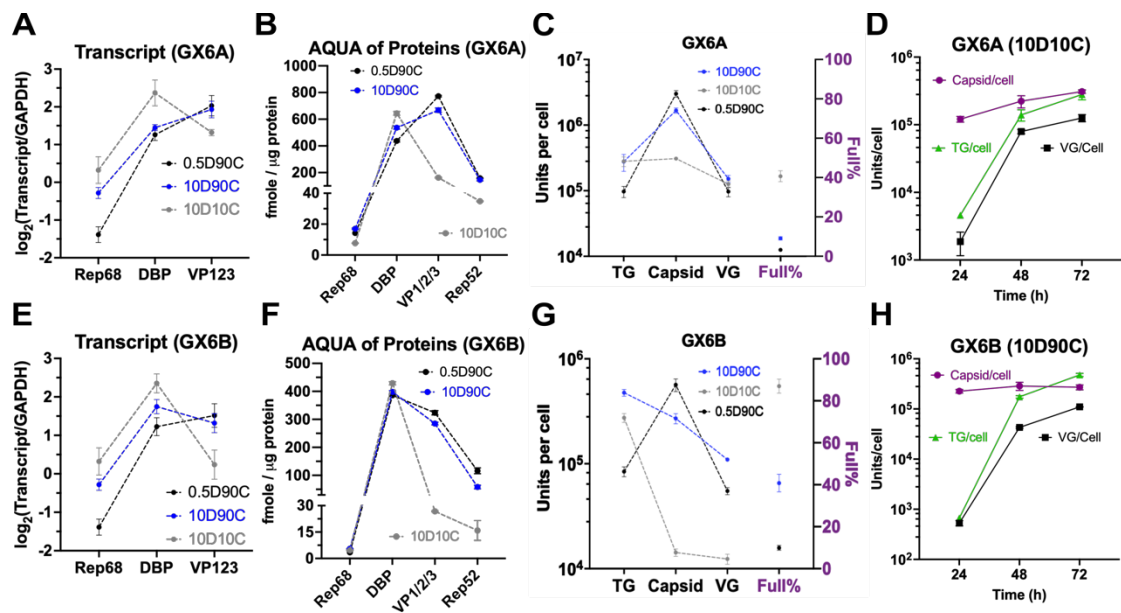


Figure 4-7. Induction conditions affected rAAV production by GX6A and GX6B cell lines. (A-D) for GX6A. (E-H) for GX6B. (A, E) Transcripts level of viral genes relative GAPDH. (B, F) Abundance of viral proteins determined by targeted proteomics. (C, G) Total intracellular virus genomes, VG and capsid titers, and full particle content. (D) Kinetic profile of capsids, total intracellular virus genome and VG in GX6A cells induced with 10D10C and (H) GX6B cells induced with 10D90C. For A and E, data represent mean and standard deviation of triplicate qPCR wells. For B and F, data represent mean and standard deviation calculated from light to heavy peak area ratios of top three fragment ions in ion chromatograms multiplied by the concentration of peptide standards. For G and H, data represent mean and standard deviation of three independent replicates.

Reducing Dox concentration from 10 to 0.5 $\mu\text{g/mL}$ reduced the Rep68 transcript level, and to a lesser degree the DBP level in GX6A cell line (Figure 4-7A). At the protein level, reducing Dox similarly affected DBP abundance in GX6A cell line (Figure 4-7B), but no reduction on Rep protein level was seen. Reducing cumate induction concentration from 90 to 10 $\mu\text{g/mL}$ lowered VP123 notably at both the transcript and protein levels (Figure 4-7A, B). Rep52 and VP123 were on the same transcript linked by IRES. At the protein level, Rep52 was also affected by the cumate concentrations but overall expressed at a much lower level (Figure 4-7B, F). In GX6B cell lines, we saw a similar effect of

reducing Dox from 10D to 0.5D on the transcript of Rep68 and DBP. But its effect at the protein level was not apparent at 72 hpi (Figure 4-7E, F).

The profoundly higher level of VP1/2/3 proteins at the high (90C) cumate level resulted in much higher levels of capsids in both GX6A and GX6B cells (Figure 4-7 C, G). The total virus genome (TG) levels responded to the reduced Dox level like the Rep68 and DBP transcripts (Figure 4-7C, G). Although AQUA assay did not capture differences at the protein level at 72 hpi, lower abundance of Rep68, DBP, and E4orf6 proteins were observed at 24 hpi (Figure 4-8D-F, Figure 4-9D-F), which might partially contribute to the difference in TG levels. However, a higher level Rep68 probably won't have a positive impact on genome amplification since stabilization of Rep68 using Shield1 resulted in a drastic decrease in rAAV titer (Figure 4-10). The level of VG (i.e., encapsidated or full particle) was affected by the levels of TG and capsid, and the packaging process in which Rep52 also plays a role. VG thus might not relate to varying induction cumate or Dox directly but rather through a more complex and indirect relationship. Nevertheless, it is clear that at a cumate induction concentration of 90 µg/mL, excess capsids resulted in high levels of empty particles. Reducing to 10 µg/mL hence resulted in higher full particle content (Figure 4-7C, G).

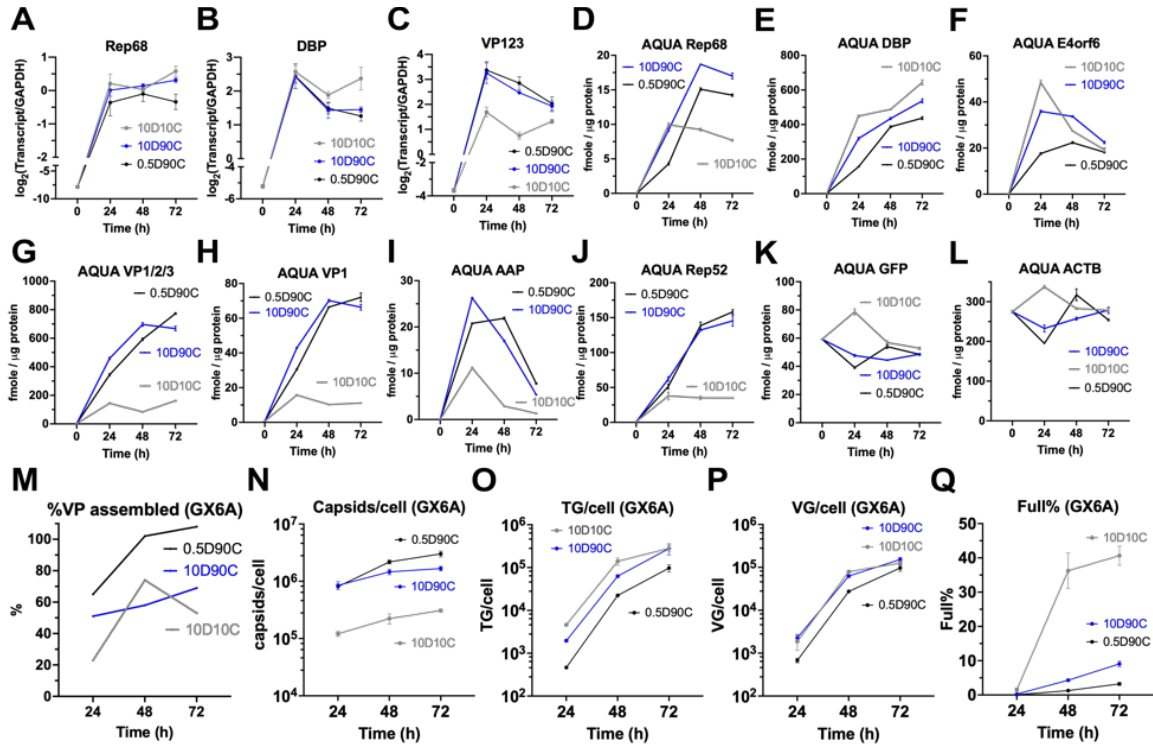


Figure 4-8. Integrated kinetic profiles of viral elements in GX6A cells.(A-C) Transcript levels of viral genes relative to GAPDH. (D-L) Absolute quantification (AQUA) of viral proteins measured by mass spectrometry with the parallel reaction monitoring (PRM) method. (M) The percentage of VP1/2/3 assembled into capsids was calculated by dividing the actual ELISA-measured capsids by the theoretical maximum number of capsids. The theoretical maximum number of capsids that could be assembled was calculated by dividing the total VP protein copies at 72 hpi by 60, assuming 60 VP1/2/3 subunits formed a capsid. (N) Capsid titers. (O) Total intracellular genomes (TG). (P) Encapsidated vector genome (VG) titers. (Q) The full particle content calculated by dividing the VG titer to the capsid titer. For A-C, data represent mean and standard deviation of triplicate qPCR wells. For D-L, data represent mean and standard deviation calculated from light to heavy peak area ratios of top three fragment ions in ion chromatograms multiplied by the concentration of peptide standards. For N-Q, data represent mean and standard deviation of three independent replicates. (I) AAP was present at lower abundance and notably decreased after reaching a peak value. (F) Interestingly, although E4orf6 expressed together with DBP via a self-cleavage peptide (P2A), E4orf6 had much lower abundance and a very different dynamic compared to (E) DBP. E4orf6/E1B55K complex participates in a wide range of activity including mRNA transport, host protein shutoff, and degradation of p53 and MRN complex to promote AAV replication.

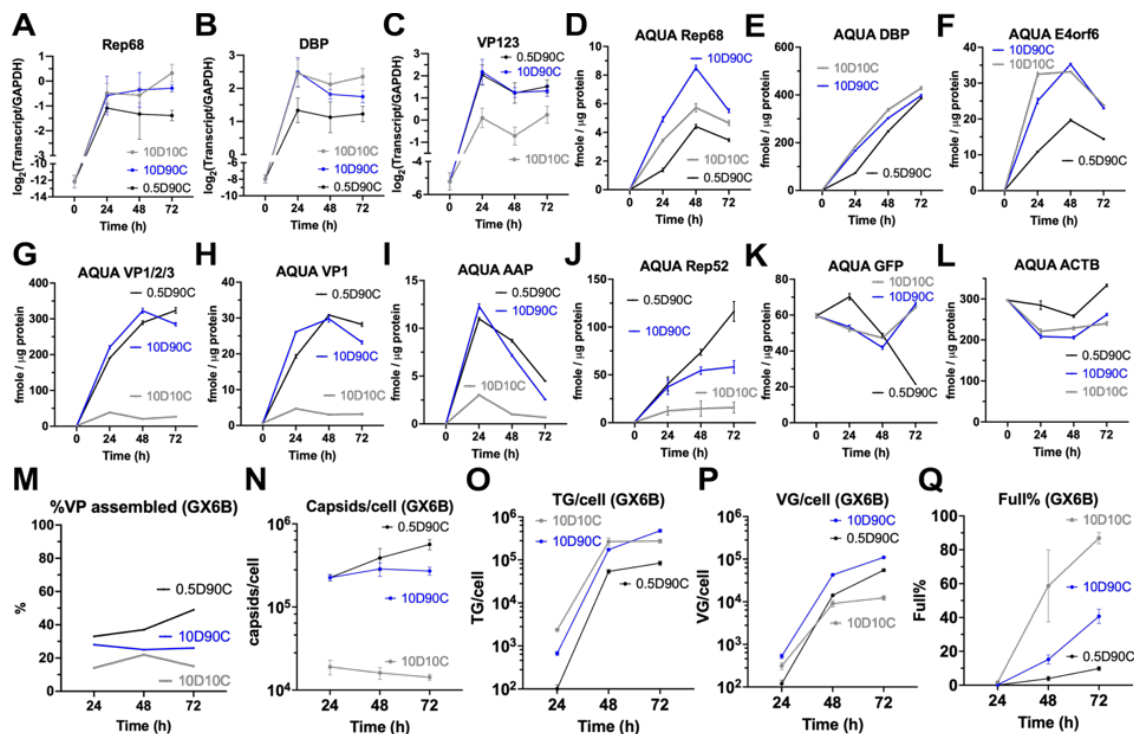


Figure 4-9. Integrated kinetic profiles of viral elements in GX6B cells.(A-C) Transcript levels of viral genes relative to GAPDH. (D-L) Absolute quantification (AQUA) of viral proteins measured by mass spectrometry with the parallel reaction monitoring (PRM) method. (M) The percentage of VP1/2/3 assembled into capsids. (N) Capsid titers. (O) Total intracellular genomes (TG). (P) Encapsidated vector genome (VG) titers. (Q) The full particle contents calculated by dividing the VG titers to the capsid titers. Mean and standard deviation were previously described.

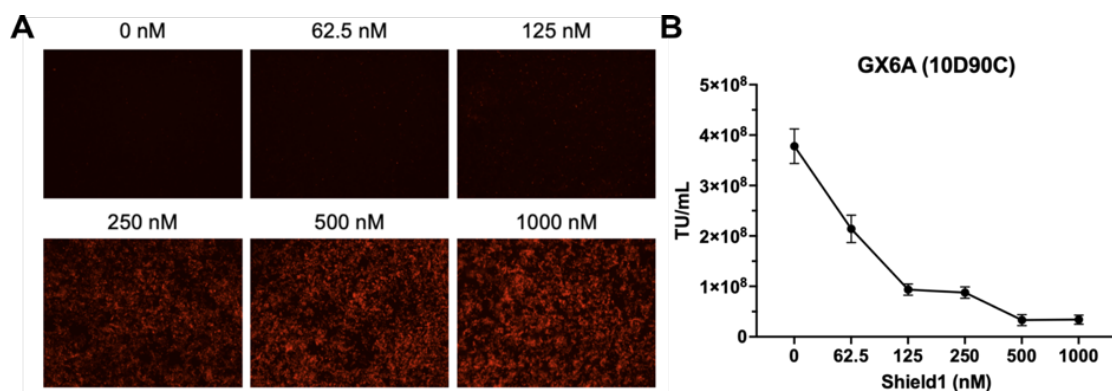


Figure 4-10. Stabilization of Rep68 with Shield1 negatively affected rAAV production.(A) mCherry fluorescence images of GX6A cells induced with 10D90C and treated with different concentrations of Shield1. (B) Dose-dependent effect of Shield1 on TU titers obtained through induction of GX6A cells.

4.4.6 Dynamics of Viral Components

What was notable about viral gene expression was their abundance levels. Compared to other viral genes, Rep68 was expressed among the lowest. However, even its transcript was on a par with that of GAPDH (Figure 4-8A). The transcript level of the most abundant VP123 was nearly ten times of GAPDH (Figure 4-8C). At the protein level, VP1/2/3 abundance was similar to or higher than that of beta-actin (Figure 4-8G, L). Also notable was that VP1/2/3 proteins rose rapidly upon induction (Figure 4-8G). At 24 hpi, it reached levels higher than beta-actin. A high percentage of VP1/2/3 proteins was assembled into capsid (Figure 4-8M, Figure 4-9M), leading to capsids rapidly reaching more than 10^5 copies at 24 hpi in both GX6A and GX6B cells (Figure 4-8N, Figure 4-9N), suggesting capsid assembly seems to be a relatively fast process.

The time profiles of capsids, TG, and VG on a per cell basis under the favorable induction condition for each cell line are shown in Figure 4-7D and Figure 4-7H. The rapid rise in assembled capsid was clearly seen. In comparison, TG rose slower and did not reach a level of 10^5 copies/cell, which was comparable to capsids/cell, until 48 hpi. Hence, for a long period of time, there were not enough TG to be packaged into capsids. For instance, a low dose of Dox (0.5D90C) resulted in slower genome amplification and a reduced TG level, which was likely the main cause of its lower VG titers and lower full particle content (Figure 4-8O, Figure 4-9O). Rep52 facilitates genome packaging. In our system, Rep52 is under the control of cumate induction. While reducing cumate level reduced excess empty particles, it also had an effect of reducing Rep52 expression (Figure 4-8J, Figure 4-9J). Here, we could not see a clear impact of Rep52 abundance on full particle content, so the effect of Rep52 abundance on genome packaging might need further investigation.

While we demonstrated that turning the inducer concentration could be used to balance a high VG titer and high full particle content, the kinetic profile of viral components pointed to using induction time profiles to achieve a balance between VG titer and full particle content. This can be achieved by delaying the induction of VP1/2/3 protein synthesis to allow genome amplification to proceed before capsids begin to accumulate.

4.5 Discussion

In this study, we demonstrated again the use of a systems synthetic biology approach to enhance the rAAV productivity in a synthetic cell line. In our earlier studies, once we established a cell line, characterized and identified factors, which limit its productivity, we devised an approach and re-constructed new producer cell lines from scratch. In this work, we took current well-producing cell lines and further enhanced them with the specific limiting element, in this case the packaging module or the *cap* gene, to boost the productivity. In the screening of clones, we used cell sorting to isolate those with more copies of integrated PM, as indicated by the intensity of the reporter gene, in spite of the possibility of an overabundance of capsids and a high representation of empty particles. When that indeed happened, we tuned induction conditions to attenuate the level of capsids, allowing for a balance between productivity and product quality. Through controlling expression levels and timing of *cap* expression, the inducible cell lines produced rAAV with high full particle content without a drop in total filled particle titers.

In this series of work, to illustrate the road to generating synthetic rAAV producer cell lines through systems synthetic biology, we have progressed from initial producer cells generating 1 VG/cell to current 10^5 VG/cell using a model cargo gene encoding GFP. Hence, we have summarized the steps we have taken for the benefit of those who will take

this synthetic route for rAAV production (Figure 4-11). Critical to our synthetic approach was the establishment of an assay cell line for efficient high producer cell isolation (Lee et al., 2023). Equally facilitating was the development of a targeted proteomics assay for quantifying viral proteins, as their roles in systems analysis were invaluable in our redesign and reconstruction cycles.

Journey to the inducible rAAV high-producing cell line

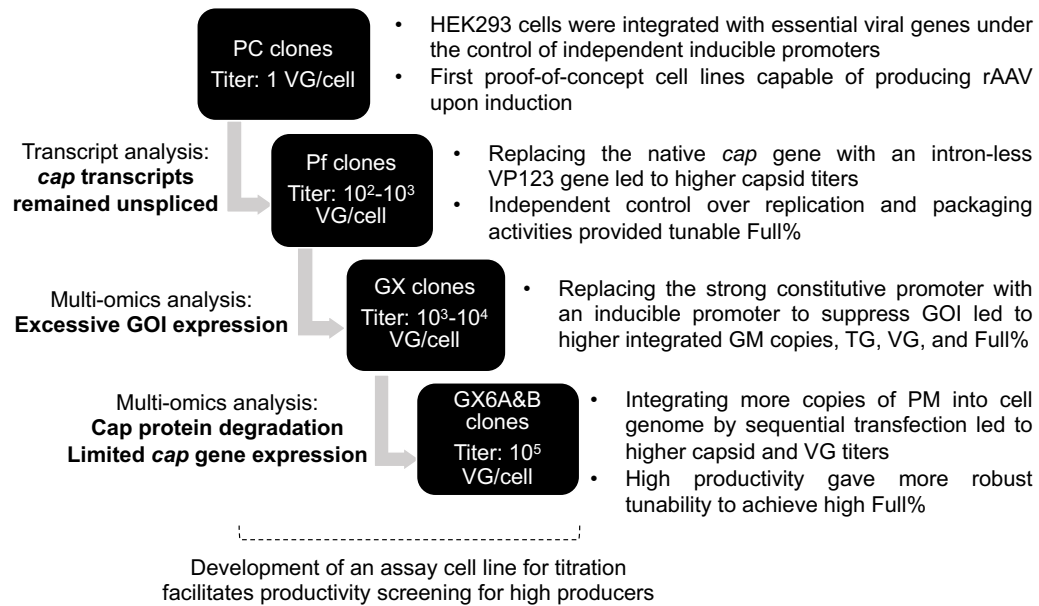


Figure 4-11. Iterative enhancement of synthetic rAAV producer cell lines. In PC cells, the spliced cap transcript level was much lower than that in the triple transfection system. We thus removed the intron and modified the translation initiation codon in the native cap gene, contributing to high capsid and VG titers in the Pf clones. Subsequently, we conducted transcriptomic and proteomic analyses of Pf3 cells, revealing extremely high GOI expression, which could potentially cause cellular burden and limit rAAV productivity. Hence, we employed another inducible promoter in GM to repress GOI expression, resulting in more integrated copies of GM and amplified virus genomes, leading to higher VG titer and full particle content in GX cell lines. In multi-omics analyses of GX cell lines, we observed Cap protein degradation and low Cap transcript and protein levels. We hence integrated more copies of PM encoding Cap proteins into GX cell lines through sequential transfection. With much higher Cap protein abundance, GX6A and GX6B produced more than 10⁵ VG per cell. The high productivity of GX6A and GX6B cell lines gave a more robust dose-dependent and temporal control over genome replication and capsid assembly to improve full particle content.

The initial PC cell lines had only relatively few copies of our genetic modules integrated into the cell genome. They produced only 1 VG per cell, most probably only a fraction of cells producing viruses, while others were not generating viral particles at all. Nonetheless, PC lines demonstrated that our modular approach and replacing viral promoters with inducible ones was a viable means of rAAV production. An important feature of the second-generation Pf cell lines was the adoption of engineered *cap* gene without the intron and with altered translation initiation to facilitate the expression of the most abundant VP3 (Urabe et al., 2002). Probably also important for the major boost of productivity was the increased copy number of integrated modules. Since we used transducing titers assessed by the assay cell line to identify and isolate high-producing clones, the resulting clones may be enriched by certain genotypic features. An important insight from the systems analysis was the potential competition for cellular resources during virus production. Several genes in the modules were expressed at very high levels, including the Cap protein and DBP, and especially GFP, which was encoded in and co-amplified with rAAV genome. The viral genome was amplified to as high as 10^5 copies/cell, equivalent to about one-sixth of the size of a host cell diploid genome. The GFP proteins in Pf3 cells were expressed to levels of most abundantly expressed cellular proteins. To synthesize multiple of these viral transcripts and proteins, and viral genomes to such high levels inevitably incurs major diversion of various cellular resources, including energy and precursor materials as well as competition for cellular machinery required for their synthesis, such as DNA and RNA polymerases, and ribosomes. At the later stage of virus production, such resource competition could become limiting. This was illustrated in the generation of GX cell lines. As the expression of VP proteins got further enhanced through

sequential transfection of the PM in those derivatives of GX lines, the possibility of cellular resources becoming limiting cannot be discounted.

While our selection of clones was largely based on TU-based productivity, equally important was the product quality, specifically full particle content. We evaluated rAAV quality after narrowing the number of clones or conditions after screenings. This approach of screening for high-titer clones and subsequently tuning the induction conditions to adjust for TG and capsid ratio was adopted in every cycle of cell line development, yielding Pf3, GX, and GX-derived cells. In all three cases, the lesson learned was that while increasing VP expression increased capsids as well as VG levels, the increase in genome packaging was lagging behind the increase in capsids, resulting in a faster rise in empty particles than encapsidated VG.

AAV has evolved intricate gene expression regulations to cooperate with helper virus. Several prior studies of AAV-Ad coinfection have described an early onset of p5-generated transcripts, which is transactivated by the Ad E1A protein, followed by transcripts driven by the p19 and p40 promoters accumulating to high levels (Mouw & Pintel, 2000; Stutika et al., 2016). At late time points, transcripts produced from p40 becomes more prevalent than p5 or p19. The AAV transcript dynamics was compounded by the explosively increasing AAV genome copy number caused by genome amplification. As large Rep proteins are not needed in very large quantity while Cap proteins are, large Rep proteins, Rep78/68, bind to the Rep binding element of the p5 promoter to repress Rep78/68 expression and transactivate the p19 and p40 promoters for small Rep protein and Cap protein production, respectively (Pereira et al., 1997). As a consequence of gene expression regulation, large Rep expression was put in check and was among the lowest

expressed viral proteins, while small Rep and Cap proteins were expressed at very high levels (Stutika et al., 2016). In our system, Rep68 was controlled by TetOn promoter. Its transcript level was among the lower ones of the viral transcript, but still at a level of only a couple fold lower than GAPDH. However, its protein level was substantially lower than other rAAV proteins. We attributed that to the linking of the destabilization domain to Rep68 protein. However, for high rAAV titer, Cap proteins must be expressed at very high levels. Unlike other systems relying on Ad infection (Chadeuf et al., 2000; Fan & Dong, 1997), HSV amplicon (Zhang et al., 1999), or SV40 origin (Inoue & Russell, 1998) to amplify the *rep/cap* copy number, the *cap* gene is not amplified in our system. Rectifying Cap protein deficiency by sequential transfection led to GX6A/B cell lines expressing over 10^5 capsid per cell with much improved productivity. This large increase in capsid formation resulted in excess empty particles as well. Nevertheless, we demonstrated that the high and early expression of *cap* gene could be tuned down and delayed through manipulation of cumate concentrations and induction timing. In line with our prior hypothesis and that of others (Destro et al., 2023), narrowing the time gap between genome amplification and capsid formation highly increased full particle content. Given that genome amplification kinetics was slow that limited rAAV packaging, it seems that the right way forward is to enhance genome amplification.

The productivity and quality of rAAV of GX6A/B cell lines are on par with those obtained through plasmid transfection and virus infection (Ou et al., 2023). For large-scale manufacturing, further increases in productivity and product quality are desired aside from adaptation to suspension culture. Given that total virus genomes only reached a level similar to or lower than that of capsids, genome amplification is a prime candidate as

probable target of intervention. Genome packaging into capsids could be controlled by several factors (Destro et al., 2023), including virus genome copy number, the packaging protein, such as Rep52 (King et al., 2001), host cell DNA replication resources (e.g., nucleotides and energy), and cellular machinery involved in the process (Weitzman & Fradet-Turcotte, 2018). With total virus genome copy in the 10^5 range and considering the time it took to reach that level, further increase in the virus genome copy number will require significant efforts. rAAV propagation is a major drain of cellular energy. In addition to ATP required for genome replication and packaging (Im & Muzyczka, 1990; King et al., 2001), wasteful synthesis of the cargo protein (Lu et al., 2023) and the potential ATP hydrolysis by DNA-independent ATPase activity of Rep68 (Schmidt et al., 2000; Zhou et al., 1999) could limit rAAV productivity. Optimizing cellular metabolism through optimization of media design and culture conditions may enhance rAAV productivity (Joiner et al., 2022).

Our synthetic cell lines included only a minimal set of AAV and helper genes. While the productivity achieved has been good, it is possible that additional helper viral genes, such as E4orf6/7 and others, which modulate host cell functions like DNA replication, cell cycle progression, and DNA repair, could aid in virus genome replication (Huang & Hearing, 1989; Polager et al., 2002). Another helper virus, herpes simplex virus (HSV), may also be a source of helper genes. For example, the catalytic activity of DNA polymerase complex formed by UL30/UL42 enhanced AAV DNA replication significantly (Alazard-Dany et al., 2009).

Like other viral infections, the replication of rAAV elicits host cell responses, including inflammatory and antiviral responses (Chung et al., 2023). The synthetic cell

lines we constructed may also elicit host responses through leaky expression of viral gene prior to induction, or after induction and expression of viral genes. Several studies have shown that knockout of apoptotic genes (Zhang et al., 2017), interferon receptors (Kim et al., 2023), or interferon-stimulated genes (Han et al., 2021) could increase virus production. It is thus conceivable that better understanding of host cell responses in synthetic cells during rAAV production will reveal potential targets for cell engineering to further enhance the productivity.

5 Multi-Omics Kinetic Analysis of Recombinant Adeno-Associated Virus Production by Plasmid Transfection of HEK293 Cells

Reproduced from Lu, M., Lee, Z., Hu, W. S. (2023) Multi-omics kinetic analysis of recombinant adeno-associated virus production by plasmid transfection of HEK293 cells. (under review)

5.1 Summary

Recombinant adeno-associated virus (rAAV) is among the most commonly used vectors for gene therapy. It is commonly produced by transfection of HEK293 cells with three plasmids each containing the vector genome including gene of interest (GOI), helper functions, and *rep* and *cap* genes for genome replication and capsid formation. To meet the potential clinical needs, the productivity of the production system needs to be enhanced. A better process characterization of the production system will further advance our insights into ways to enhance the productivity. Here, we employed transcriptomic analysis to quantify the dynamics of different isoforms of viral transcripts and to assess the shift of cellular physiology, and deployed targeted proteomic analysis for absolute quantification of viral proteins and tandem mass tags (TMT) for assessing cellular responses at the protein level. Functional analysis at transcriptome and proteome levels identified defense and immune response, unfolded protein response, p53 signaling as enriched. The small molecule additive intervention study based on functional analysis showed the potential of such omics-guided productivity enhancement. Together, multi-omics analysis advanced understanding of rAAV production and provided insight into enhancing rAAV production by plasmid transfection.

5.2 Introduction

Recombinant adeno-associated virus (rAAV) is one of the most commonly used gene delivery vehicles for human gene therapy owing to its tissue tropism, safety and long-term persistency (Samulski & Muzyczka, 2014). Several rAAV-based gene therapies, including Luxturna for retinal dystrophy, Zolgensma for spinal muscular atrophy, Hemgenix for hemophilia B, and Elevidys for Duchenne muscular dystrophy have been approved by FDA and reached clinical applications (He et al., 2021; Herzog et al., 2023). Generally, except for a small organ like eyes, a very large dose of rAAV is required in each treatment (e.g. around 10^{14} vector genome/kg) (Mendell et al., 2017), and the manufacturing of large quantities of rAAV for clinical applications poses a significant challenge to current production platforms.

AAV is a small single-stranded DNA capsid virus. Its small genome essentially comprises two inverted terminal repeats (ITRs) flanking *rep* and *cap* genes (Naso et al., 2017). AAV *rep* and *cap* genes encode large (Rep78/68) and small (Rep52/40) Rep proteins, and three capsid proteins (VP1, VP2, VP3) and assembly-activating protein (AAP), respectively, through alternative splicing and multiple translation initiation sites. Cap gene has many serotypes which gives different species- and tissue-specificity (Mingozzi & High, 2013). AAV is a replication-defective virus and requires coinfection of a helper virus, such as Adenovirus (Ad) or Herpes Simplex virus (HSV), to provide helper factors to activate *rep/cap* gene expression and replicate its genome (Merten, 2016). To generate rAAV, the *rep* and *cap* genes in the AAV genome are replaced by the coding sequence of the gene of interest (GOI) so that the genome of the virus produced lacks *rep* and *cap* genes and cannot replicate in target cells. In addition to the rAAV genome, one

thus needs to supply *rep*, *cap*, and helper genes in *trans*. Upon subsequent transduction to target cells, the rAAV expresses GOI but cannot replicate.

rAAV is commonly produced by transient transfection of HEK293 cells with multiple plasmids, one encoding the therapeutic GOI flanked by AAV ITRs, the others containing AAV *rep* and *cap* genes, and adenoviral helper genes *E2A*, *E4*, and *VA RNA* (Clément & Grieger, 2016). Additionally, HEK293 cells harbor another Ad helper gene E1. Different approaches have been taken to increase rAAV productivity of plasmid transfection, including optimization of plasmid design (Emmerling et al., 2016), transfection process (Zhao et al., 2020), and treatment with small molecule chemicals (Scarrott et al., 2023). A number of alternative production systems use viral vectors based on Baculovirus or HSV (Eduard Ayuso et al., 2010). Each system employs two recombinant viral vectors, one carries rAAV genome, the other encodes Rep and Cap proteins. Co-infection of the two viral vectors into insect or human host cells leads to the production of rAAV. HeLa cells have been stably integrated with AAV *rep* and *cap* genes, upon co-infection of a hybrid Ad virus containing rAAV genome and a wild type Ad, rAAV is generated (Chadeuf et al., 2000). These systems require the preparation of viral vectors to seed the production of rAAV (Wright, 2008) as well as post-production removal of contaminating viruses (Eduard Ayuso et al., 2010). rAAV produced in insect cells also pose additional concerns about different post-translational modifications as compared to human cell produced product (Rumachik et al., 2020). Recently, a synthetic biology-based system was developed by integrating rAAV genome, essential AAV and Ad helper genes, under the control of inducible promoters, into HEK293 genome (Lee et al., 2022). Upon

induction, the system produced rAAV. To date, multi-plasmid transfection-based production method is still the most commonly used.

Even though multi-plasmid transfection system is commonly used in rAAV production for gene therapy, there has not been many reports on the dynamics of various viral components and the rAAV. Since viruses are parasitic in nature and rely on host cell materials and machinery for their replication, the dynamics of various viral components, assembly and packaging are highly dependent on host cell's responses to their replication. On the other hand, it has been shown that the severity of host antiviral response could affect the productivity of virus (Samuel, 2001; Ye et al., 2021). Kinetic studies on rAAV production hence need to be global, at both virus and host cell levels. To that end, a proteomic study using label-free quantification on host cell proteins showed altered proliferation and metabolism upon rAAV production using triple plasmid transfection (Strasser et al., 2021). RNAseq studies reported transcriptional upregulation of inflammatory and antiviral response during rAAV production by plasmid transfection, again on host cell genes (Chung et al., 2023; Wang et al., 2023).

This work examined the kinetics of rAAV production in plasmid-transfected HEK293 cells by transcriptomic analysis and quantitative proteomics to enable a thorough profiling of viral gene expression and global functional alterations at the host cell level (Figure 5-1). The absolute quantification (AQUA) through targeted quantitative proteomics allowed for a first quantitative look at the kinetics of all major viral components, while tandem mass tags (TMT) based proteomic lent us corroboration of host functional response at the protein level (Vogel & Marcotte, 2012). Based on the findings of functional analysis, we showed that treatment of the plasmid transfection-based production system

with small molecules could enhance rAAV productivity. The work hence demonstrated the potential impact of such integrated multi-omics study on viral vector production.

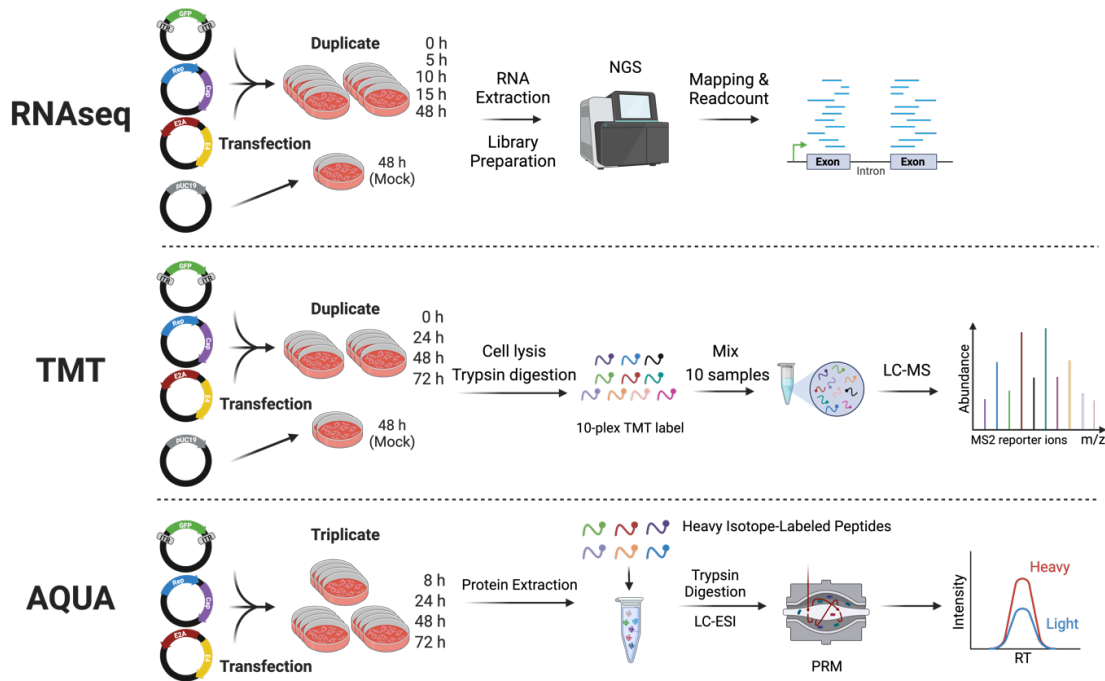


Figure 5-1. The scheme of multi-omics analysis of HEK293 cells during rAAV production through plasmid transfection. RNAseq and TMT-based proteomics were used to characterize both viral and cellular gene expression at transcript and protein levels respectively. Absolute quantification (AQUA) of viral proteins by targeted proteomics using parallel reaction monitoring (PRM) method with heavy-isotope labeled peptides was performed. Two independent replicates at each timepoint were analyzed for transcriptomic and TMT-based proteomic analysis. For AQUA, three independent replicates at each timepoint were analyzed.

5.3 Materials and Methods

5.3.1 Cell Culture and rAAV Production through Triple Transfection

HEK293 cells (Cell Biolabs, Inc.) were cultured in Dulbecco's Modified Eagle Medium (Gibco) supplemented with 10% Fetal Bovine Serum (Gibco) and antibiotic/antimycotic (Gibco) in T-flask. Triple transfection with pAAV-CAG-EGFP (Addgene #37825), pHelper and pAAV-DJ (Cell Biolabs, Inc.) (Molar ratio = 1:1:1) was conducted on HEK293 cells using Transporter 5 transfection reagent (Polysciences)

(DNA:PEI = 1:4 [w/w]). Cells were harvested at timepoints of interest post-transfection for acquisition of crude rAAV lysate, intracellular DNA, RNA, and proteins.

5.3.2 rAAV Preparation and Characterization

Preparation of crude rAAV lysate, qPCR assay of vector genome titration, and ELISA were previously described (Lee et al., 2022). Total intracellular virus genome (TG) copies, including free virus genomes and encapsidated vector genomes, were measured by qPCR using total intracellular DNA.

5.3.3 Illumina Sequencing and Differential Expression Analysis

RNA was extracted using the RNeasy Mini kit (Qiagen). Library preparation, read length and depth, sequencing platform, read mapping and counting, and TPM calculation were previously described (Lee et al., 2022). Differentially expressed genes between different conditions were identified, and fold-changes were determined using a quasi-likelihood negative binomial generalized log-linear model in the EdgeR package in R (Robinson et al., 2009). Genes with a CPM (count per million) smaller than 3 in all samples were excluded from further analysis.

5.3.4 Quantitative Proteomics Analysis using Tandem Mass Tags

Cell pellets were reconstituted in lysis buffer, then denatured and reduced. Protein samples were quantified by Bradford assay, then digested into peptides using trypsin. TMT labeling, data acquisition by Orbitrap Fusion (Thermo Fisher Scientific), and database searching for protein detection in Proteome Discover (Thermo Fisher Scientific) were previously described in detail (Yang et al., 2022).

5.3.5 Targeted Quantitative Proteomics Analysis

Sample preparation, data acquisition by Orbitrap Fusion, and data analysis in Skyline (MacLean et al., 2010) were previously described in detail (Lee et al., 2022). Light-to-heavy peptide ratios were calculated with peak areas, and then ratios were used to determine endogenous protein concentrations. Heavy isotope-labeled peptides used in this study are shown in Table 5-1.

Table 5-1. Sequences of heavy isotope-labeled peptides used in the targeted proteomics assay for absolute quantification (AQUA).

AQUA Peptide Sequence	Targeted Protein	Coding Transcript
GIEPTLPNWFVAVT[K_C13N15]	Rep78/68	Rep78/68
NVSLPVAHSDA[R_C13N15]	DBP	E2A
HPPPQILI[K_C13N15]	VP1, VP2, VP3	VP1, VP2/3A
APTEWVIP[R_C13N15]	AAP	VP2/3A
QLDSGDNPYL[K_C13N15]	VP1	VP1
FSVSGEGEGDATYG[K_C13N15]	EGFP	EGFP
TAPDYLVGQQPVEDISSN[R_C13N15]	Rep78/68/52/40	Rep78/68/52/40

5.3.6 Functional Class Analysis

Overrepresentation analysis (ORA) and functional annotation clustering were carried out using DAVID bioinformatics resources (Huang et al., 2009). In RNAseq analysis, a more stringent significance threshold FDR $p < 0.05$ and > 1.5 fold-change was used for using 0 hpt as control while a relaxed threshold FDR $p < 0.05$ was used for using Mock as control. In proteomic analysis, a relaxed significance threshold nominal $p < 0.05$ was used for all comparisons without a fold-change criterion due to a smaller dynamic range of protein abundance to have at least more than 500 differentially expressed proteins. Lists of differentially expressed genes (DEGs) were tested against all identified transcripts

or proteins as background, and Gene Ontology (GO) (Ashburner et al., 2000) and KEGG pathway (Kanehisa & Goto, 2000) databases were used for functional annotation. Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) using all expression data as input were also conducted with the default setting and GO and KEGG databases.

5.4 Results and Discussion

5.4.1 Summary of Data

HEK293 cells were transfected with equal molar quantity of three plasmids containing rAAV genome with EGFP as GOI, AAV *rep-cap* genes, and Ad2 helper genes (*E2A*, *E4*, and the non-coding *VA RNA*). Cell samples of transfected cells were taken at different timepoints for RNAseq and proteomic analysis. Untransfected HEK293 cells, and a Mock control, in which cells were transfected by an equal mass quantity of pUC19 plasmid devoid of transgene, were included as controls.

Principal component analysis was performed on duplicate samples of each condition, and expected similarity was seen (Figure 5-2). For RNAseq, around 20,000 transcripts were identified in all samples. Pairwise comparison of 5 vs. 0 hour post transfection (hpt), 10 vs. 0 hpt, 15 vs. 0 hpt, 48 vs. 0 hpt, Mock vs. 0 hpt (fold-change>1.5, FDR<0.05), and 48 hpt vs. Mock (FDR<0.05) were performed to identify differentially expressed genes (DEGs). The number of DEGs increased progressively as expected, and fewer DEGs were seen in 48 hpt vs. Mock (Figure 5-3A-F). For TMT-based proteomics, around 5,000 proteins were detected and quantified with at least two peptides in all samples. Pairwise comparison of 24 vs. 0 hpt, 48 vs. 0 hpt, 72 vs. 0 hpt, Mock vs. 0 hpt, and 48 hpt vs. Mock were performed to identify differentially expressed proteins (nominal p -val<0.05), and around 500 differentially expressed proteins were seen in each comparison

(Figure 5-4A-E). Differentially expressed transcripts with notably high fold-changes in 48 vs. 0 hpt and 48 hpt vs. Mock were shown in Figure 5-3G and H, and proteins were shown in Figure 5-4F and G.

The depth of coverage was much lower in proteome than transcriptome as expected. It is also not surprising that highly abundant transcripts had a higher probability that their proteins were also identified in TMT proteomics (Figure 5-5).

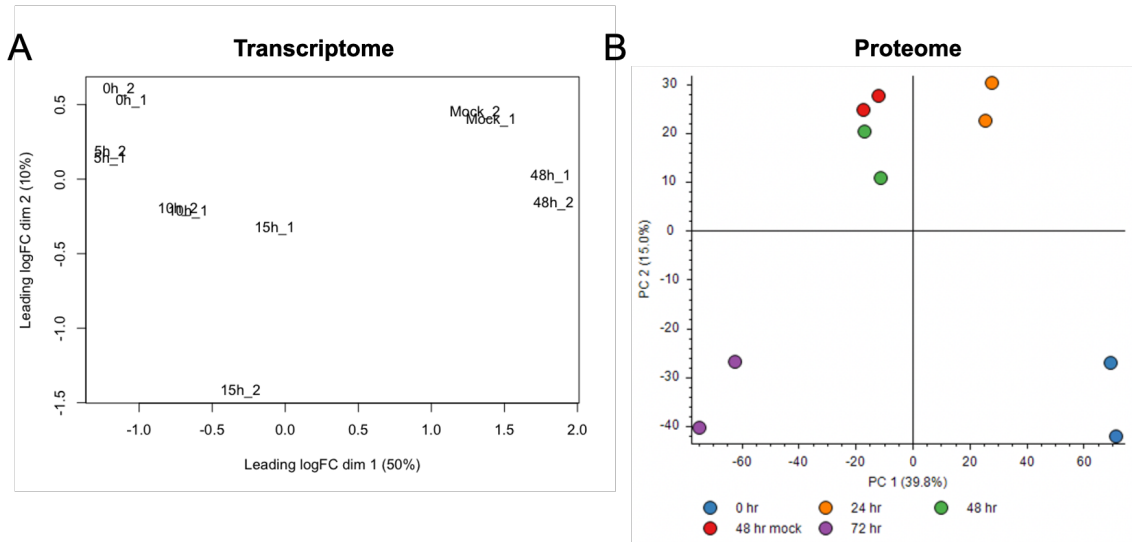


Figure 5-2. Principal component analysis (PCA) of samples of (A) RNA-seq and (B) TMT-based proteomics. PCA analysis was performed using EdgeR and Proteome Discoverer for RNAseq and TMT-based proteomics respectively. Duplicate samples of each condition had high similarity, and 48 hpt and Mock samples were close to each other.

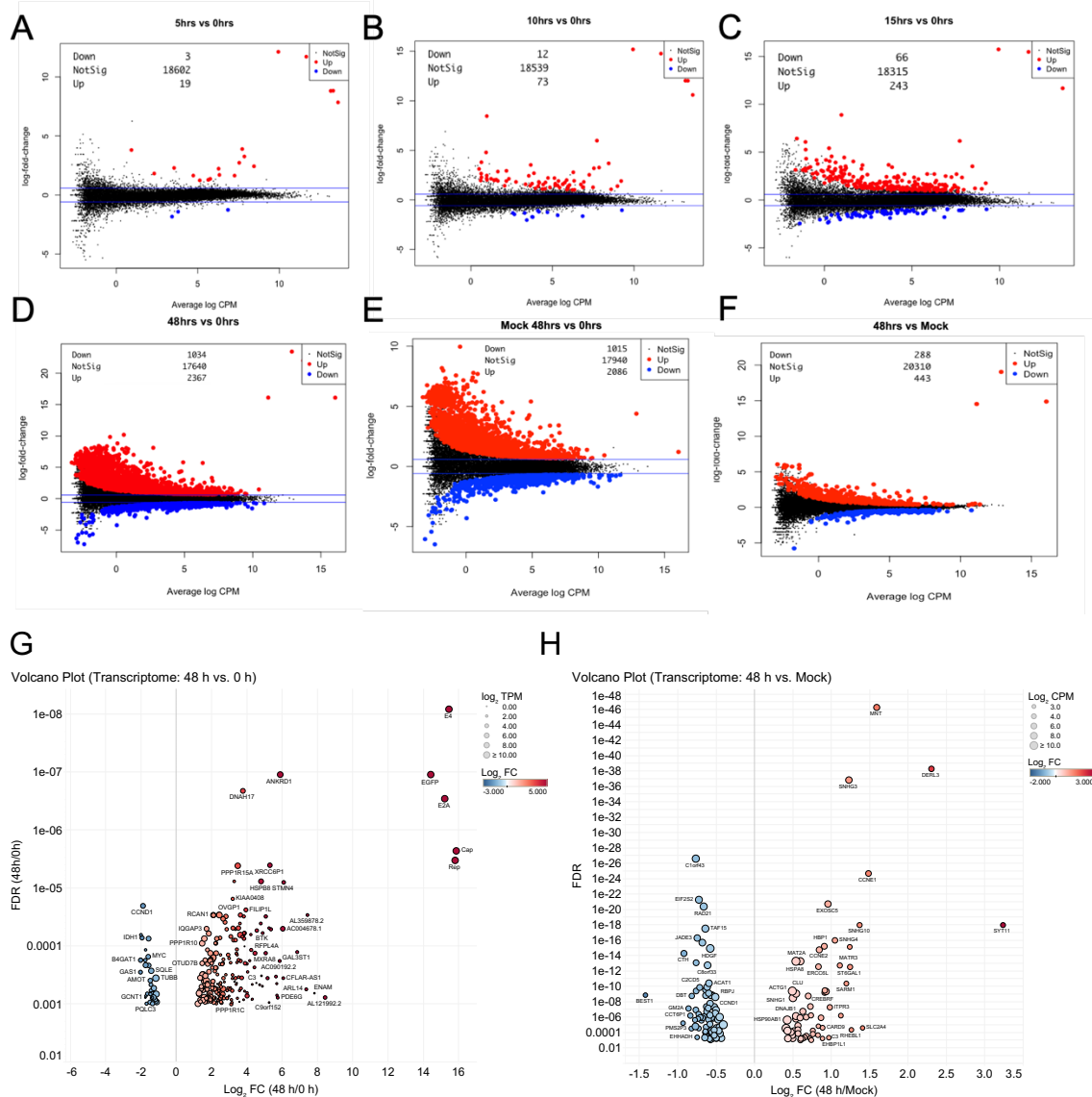


Figure 5-3. Differential expression analysis of RNA-seq data using EdgeR. MA plots and numbers of differentially expressed transcripts of different comparisons are shown: (A) 5 vs. 0 hpt (B) 10 vs. 0 hpt (C) 15 vs. 0 hpt (D) 48 vs. 0 hpt (E) Mock vs. 0 hpt and (F) 48 hpt vs. Mock. Differentially expressed transcripts with notably high fold-changes in (G) 48 vs. 0 hpt and (H) 48 hpt vs. Mock are labeled with gene name. Significance threshold adjusted p-value < 0.05 and fold-change > 1.5 was used for 5, 10, 15, 48 vs. 0hpt and Mock vs. 0 hpt. Significance threshold adjusted p-value < 0.05 was used for 48 hpt vs. Mock. In RNA-seq result, 21378 (0 hpt), 21511 (5 hpt), 21427 (10 hpt), 21857 (15 hpt), 25108 (48 hpt), 25240 (Mock), and 18384 transcripts in intersections were identified. Fewer DEGs were identified at early timepoints while more than 2000 genes were differentially expressed at 48 hpt. Fewer DEGs were identified in 48 hpt vs. Mock comparison.

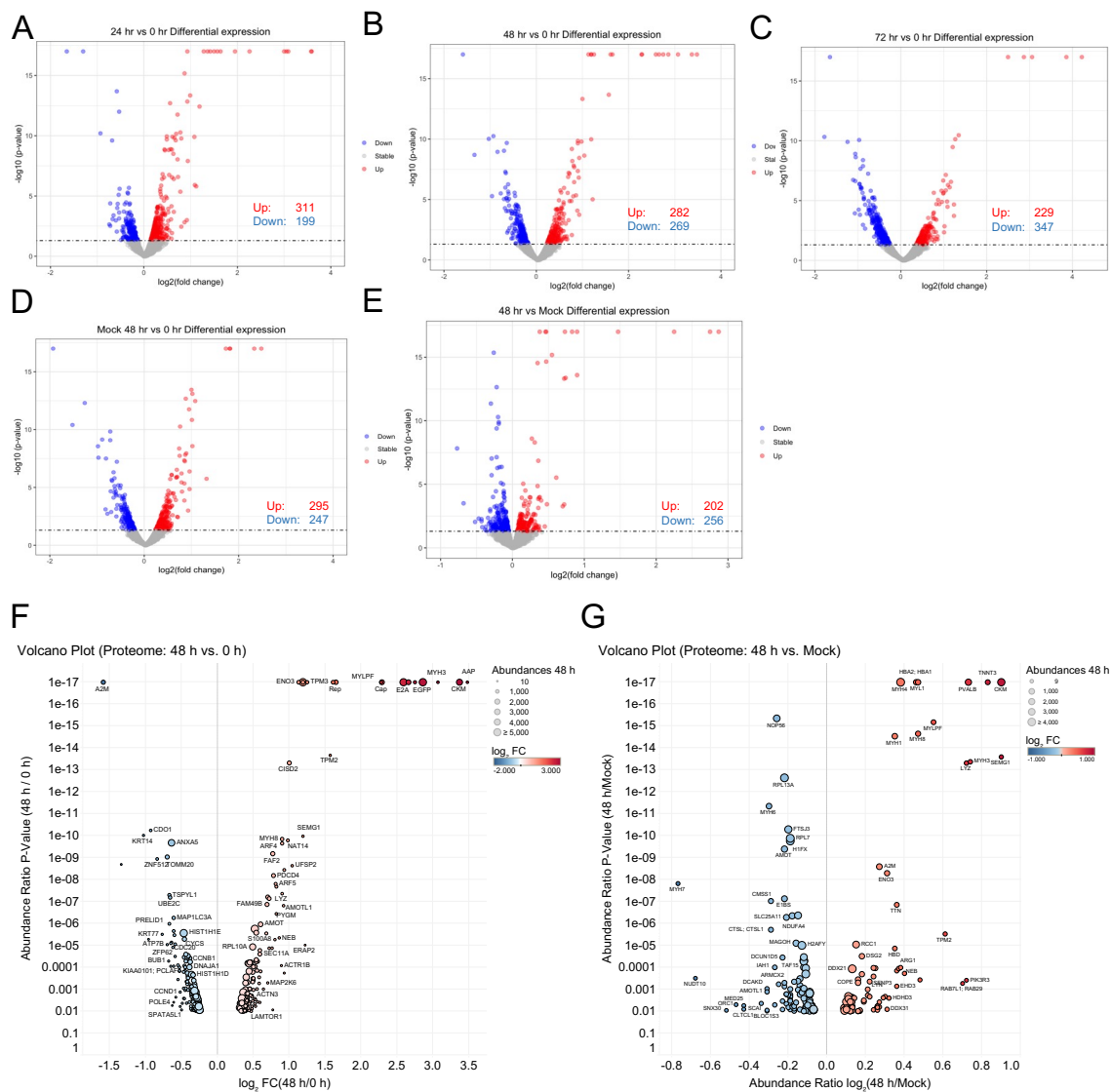


Figure 5-4. Differential expression analysis of TMT data using Proteome Discoverer. Volcano plots and numbers of differentially expressed proteins of different comparisons are shown: (A) 24 vs. 0 hpt (B) 48 vs. 0 hpt (C) 72 vs. 0 hpt (D) Mock vs. 0 hpt (E) 48 hpt vs. Mock. Differentially expressed proteins with notably high fold-changes in (F) 48 vs. 0 hpt and (G) 48 hpt vs. Mock are labeled with gene name. Significance threshold nominal p-value < 0.05 was used for all comparisons. In TMT assay result, 6265 proteins were identified, and 5461 proteins with at least two identified peptides were quantified. The degree of protein differential expression increased progressively with hpt as expected, and around 500 proteins were differentially expressed in each comparison. The p-values were calculated by using the background population of ratios for all peptides and proteins to determine whether any given a single peptide or protein is significantly changing relative to that background. Since the dynamics range of proteins was much smaller than transcripts, fold-change criterion was not applied.

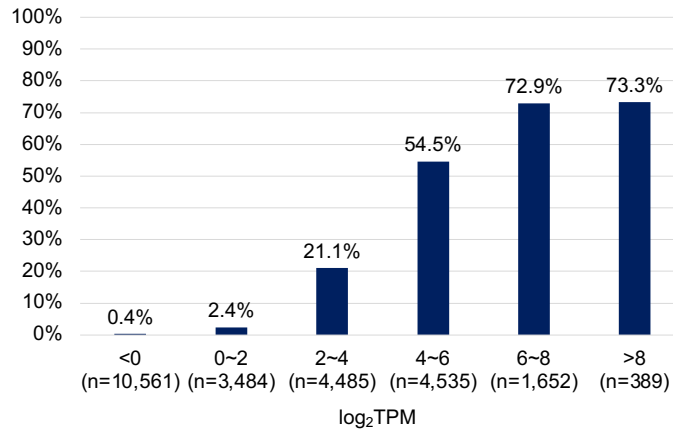


Figure 5-5. The percentage of transcripts whose encoding proteins were identified in TMT-based proteomics. Transcripts were first grouped based on their TPM, then percentage was calculated by dividing number of proteins identified in TMT-based proteomics by number of total transcripts in that group.

5.4.2 Dynamics of Viral Transcripts and Proteins upon Triple Transfection

5.4.2.1 Dynamics of viral transcripts

The abundance and dynamics of viral transcripts and proteins are shown in Figure 5-6A-E. *Rep* and *cap* genes of AAV are transcribed into multiple transcripts from three promoters p5, p19, and p40 and translated into at least 8 proteins (Figure 5-7A). All viral transcripts initially increased rapidly, by 5 hpt their abundance level spanned over nearly two orders of magnitude with EGFP and transcripts from the helper plasmid (Figure 5-6A, Figure 5-7D, E) increased most prominently. Triple transfection differs from AAV infection and replication in a major way: in AAV infection, the genome is being amplified, as well as the gene copies of *rep* and *cap*. In triple transfection, only EGFP-encoding rAAV genome is amplified, but not *rep* and *cap* genes. In fact, barring from plasmid degradation, the copy number of helper, *rep* and *cap* genes would be constant. The three Cap proteins are present in a capsid at a proportion of VP1:VP2:VP3=1:1:10. VP23A (VP2, VP3, AAP coding transcript) was synthesized about 10 times higher than VP1. In addition to generated

VP1 and VP23A transcript, p40 also generated an unspliced variant. Interestingly, the level of unspliced variant was even higher than VP23A. In our recent work, on a rAAV producing synthetic cell line, the intron was removed to eliminate the unspliced variant, and the productivity was improved (Lee et al., 2022). Rep78/68 and Rep52/40 transcripts were transcribed somewhat lower; however, their TPM value in the thousand range would rank them as highly abundant among cellular genes. Overall, p40 promoter drove total VP transcripts to higher levels than p19 for Rep52/Rep40, which in turn was transcriptionally more active than p5 for Rep78/68. Similar observation on transcript abundance has been reported before for wtAAV/Ad coinfection (Mouw & Pintel, 2000; Stutika et al., 2016). This may not be surprising as the *rep-cap* plasmid used in transfection retained the wtAAV regulatory elements in its genome.

The two Ad helper protein genes, E2A and E4 encoded in pHelper plasmid and two others, E1A and E1B, integrated in HEK293 genome all have alternatively spliced isoforms with varying degree of abundance (Figure 5-7B-E). Their expression increased rapidly right after transfection and stayed at very high levels. The early rise in E2 and E4 helper reflected their early gene nature. E1A and E1B genes which were integrated in HEK293 genome, expressed at moderately high levels even before transfection and increased by several folds after transfection.

In triple transfection, a large number of each plasmid would likely be taken up by a cell, entering nucleus and undergoing transcription. However, only the rAAV genome is amplified. Hence, the EGFP transcripts were transcribed from both transfected plasmids and amplified rAAV genome. Due to the high copy number of rAAV genome, the

transcript of EGFP increased to over 5% of total cellular mRNA as shown in Figure 5-6A and Figure 5-7F.

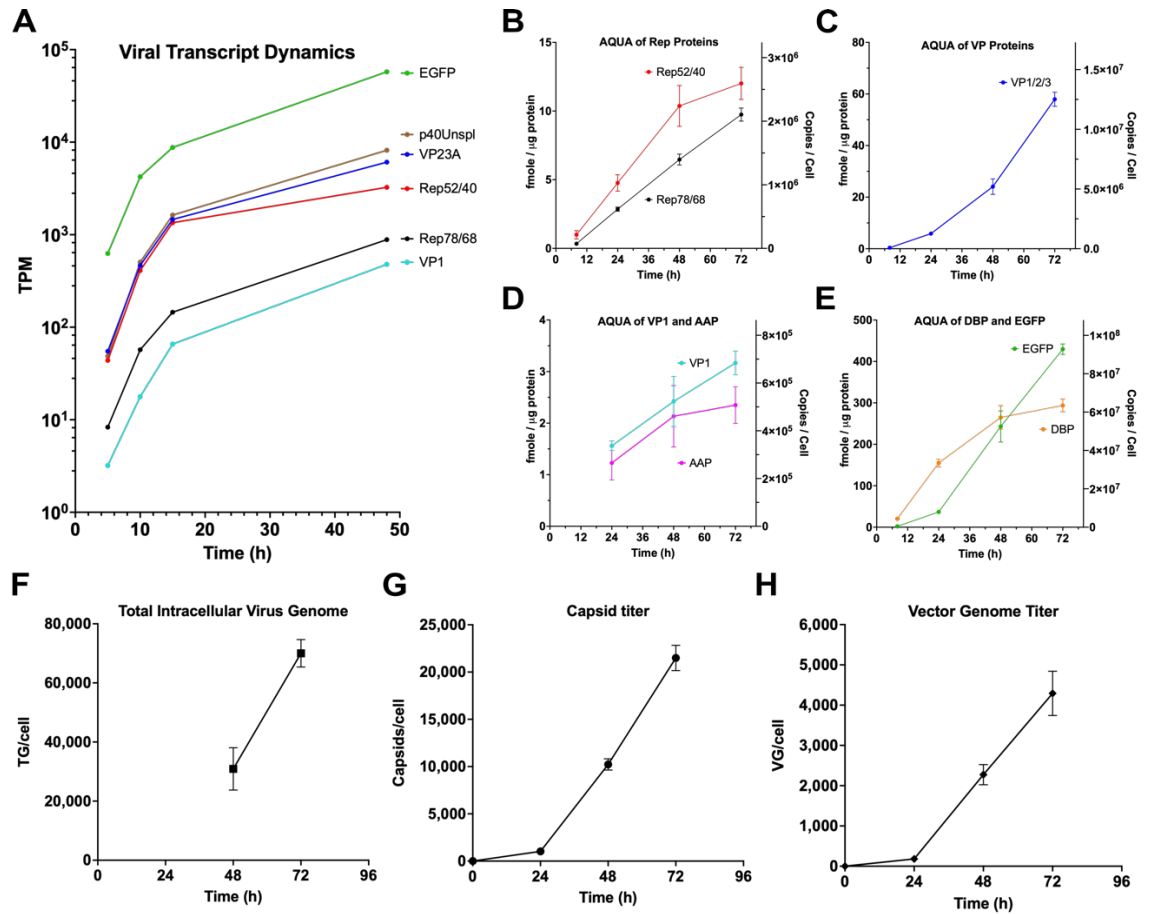


Figure 5-6. Viral transcript and protein dynamics.(A) TPM of viral transcripts determined using RNAseq data. AQUA of (B) Rep, (C) VP1/2/3, (D) VP1, AAP, (E) DBP and GFP by LC-MS using PRM method with heavy isotope-labeled peptides. Protein copies per cell were estimated by assuming 360 pg protein per HEK293 cell. AQUA measured the total VP proteins including those in free form and in assembled capsid particles. (F) The total intracellular virus genome (TG) copies per cell quantified by qPCR and referenced to dummy triple transfection and the DNA sequence with known copy number in cells. (G) rAAV2 capsid titer measured by ELISA to approximately estimated copies of assembled capsids produced from triple transfection with pAAV-2/DJ. (H) Benzonase-resistant encapsidated vector genome (VG) copies per cell measured by qPCR. For B-H, data represent mean and standard deviation of three independent replicates.

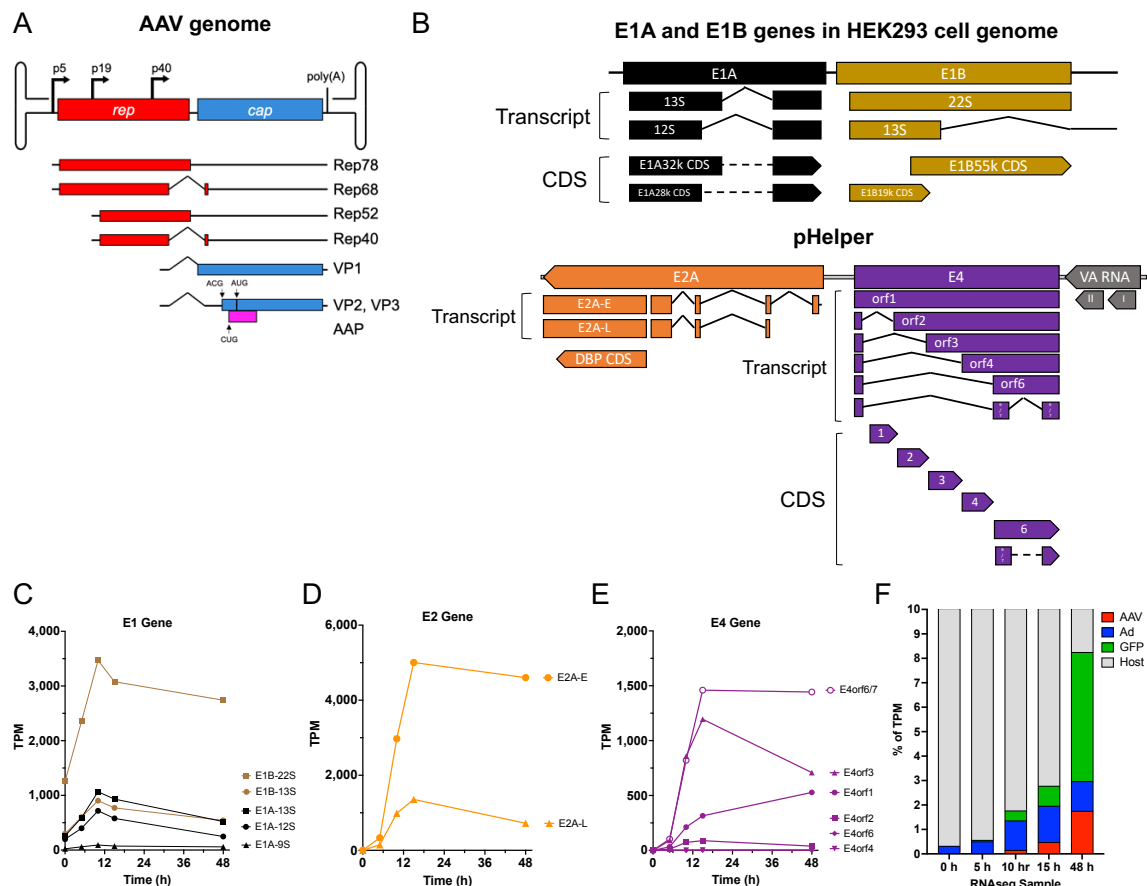


Figure 5-7. Schematics of viral genome and dynamics of helper transcripts. (A) AAV genome structure (B) Transcripts and protein coding sequences of adenoviral helper E1, E2, and E4 genes. Transcript abundance of (C) E1, (D) E2, and (E) E4 alternatively spliced transcripts. E1A32K protein produced from E1A-13S induces p53-dependent apoptosis, while E1B19K produced from E1B13S and E1B55K produced from E1B22S inhibit apoptosis. In addition, E1B55K complexed with E4orf6 also induces degradation of MRN complex, which inhibits AAV genome replication. Non-coding VA RNA inhibits PKR-mediated anti-viral response to increase viral protein synthesis. E2A-E transcript encodes DNA binding protein (DBP) which facilitate AAV genome replication by increasing DNA polymerase processivity. (F) Percentage of TPM of transcripts originally from AAV, Ad, GFP, or host genes. Aside from constitutively expressed E1A and E1B, Ad viral transcripts increased very fast and accounted for 1.5% of total transcripts at the early stage. In contrast, AAV transcripts increased to higher levels and accounted for 2% of total transcripts at 48 hpt. Because of genome amplification, GFP transcripts produced from rAAV genome increased dramatically at 48 hpt and accounted for 5% of total transcripts.

5.4.2.2 Dynamics of viral proteins

Expression of viral proteins including AAP, produced from a different ORF of the same transcript of VP2 and VP3 (i.e., VP23A), was quantified using heavy isotope-labeled

peptides (Figure 5-6B-E). Targeted quantitative proteomics with spiked peptide standards allowed for AQUA in femtomole per microgram protein or copies per cell. The peak values of abundance level of different proteins spanned over 100-fold. Protein dynamics assessed by TMT showed similar trends as in AQUA data (Figure 5-8A-G). Unexpectedly, abundance of constitutively expressed E1 proteins also changed somewhat upon transfection (Figure 5-8F).

Helper protein, DNA binding protein (DBP), was the most abundant viral proteins, reaching over 6×10^7 molecules/cell (Figure 5-6E). The next was total VP subunits at $\sim 1.3 \times 10^7$ molecules/cell and VP1 at $\sim 7 \times 10^5$ /cell (Figure 5-6C, D). Although VP2 subunit could not be uniquely quantified, the ratio of VP1 and VP2/3 was around 1:17 which was close to the consensus VP ratios (Gonçalves, 2005). Small Rep proteins was present at a slightly higher level than large Rep proteins (Figure 5-6B).

EGFP reached almost 10^8 copies per cell, a consequence of the extremely high transcript level driven by genome amplification (Figure 5-6E). Such high-level transcription and translation of a “wasteful” product inevitably diverted cellular resources like raw materials and machinery away from rAAV production as have been observed previously (Strobel et al., 2015).

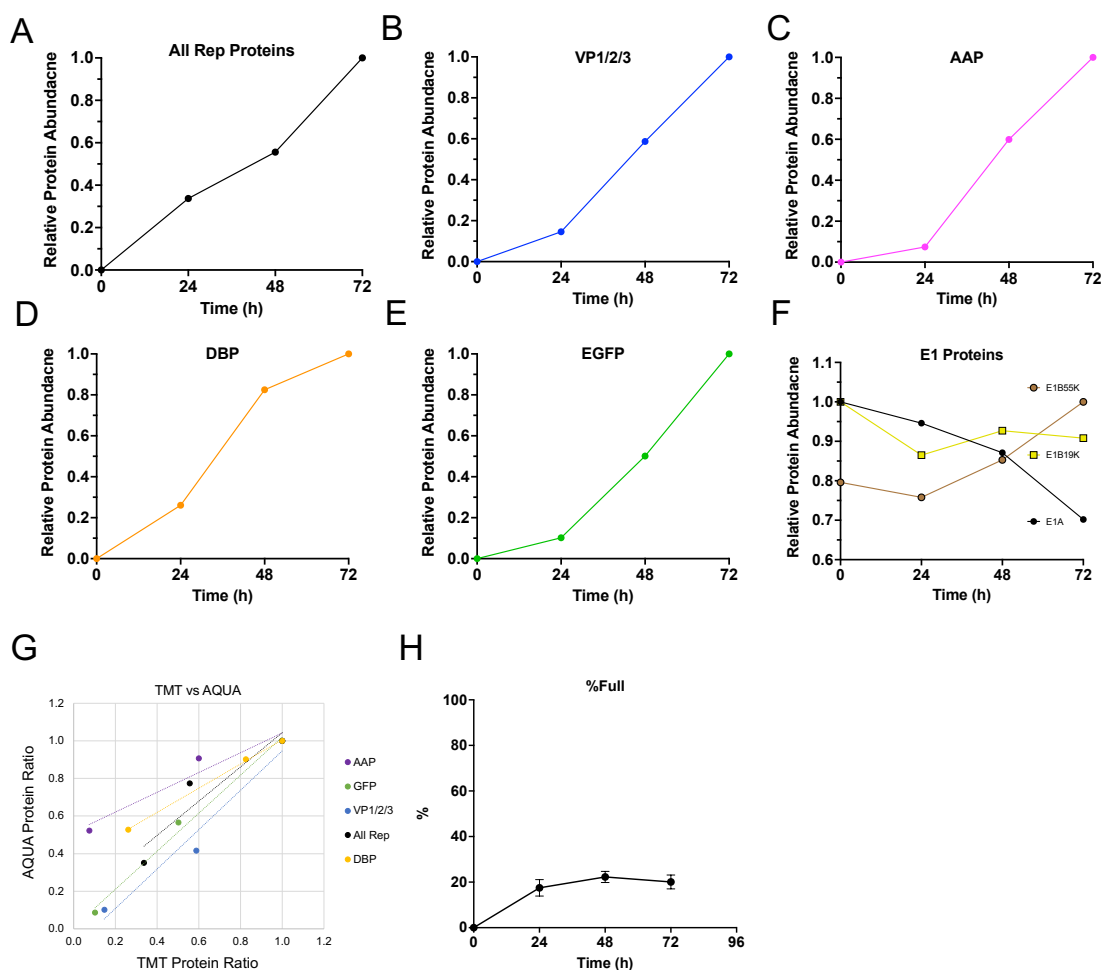


Figure 5-8. Protein dynamics determined by TMT-based proteomics analysis. (A) Rep, (B) VP1/2/3, (C) AAP, (D) DBP, (E) GFP and (F) E1. (G) Comparison of relative protein quantification by TMT-based and targeted proteomics. The relative protein ratios were calculated by normalized protein abundance at 24 hpt, 48 hpt, and 72 hpt to protein abundance at 72 hpt. Overall, the result from two assays is consistent, but AAP seems to be somewhat deviated. This might result from the lower resolution and sensitivity of TMT-based proteomics, especially for a low abundance protein like AAP. (H) Full particle contents calculated by dividing VG by capsid titers.

5.4.2.3 Kinetics of viral particle production

We quantified the total rAAV genomes in the cell (TG) (Figure 5-6F), total capsids (Figure 5-6G), as well as the genomes packaged into capsids, i.e., encapsidated genome or vector genome (VG) (Figure 5-6H). We assumed that equal number of total capsids and total genomes were packaged together to form full particles, leaving behind excess

genomes and unpackaged capsids as empty ones. At 72 hpt, on a per cell basis, there were around 7×10^4 TG, 4×10^3 VG, 1.25×10^7 VP1/2/3 proteins and 2×10^4 capsids. Assuming 60 copies VP1/2/3 form a capsid, those assembled capsids would have consumed around 1.2×10^6 capsids while leaving behind about 90% of unassembled VP1/2/3. The genome packaged into assembled capsids was inefficient, out of 7×10^4 TG only 4×10^3 were encapsidated to become VG ending up only 20% of assembled capsids contained rAAV genome (Figure 5-8H). The data suggest that genome encapsidation is likely a rate-limiting step for rAAV production and improvement in genome encapsidation efficiency is called for.

5.4.3 Host Cell Transcriptome and Proteome Changes Following Triple Transfection

Differentially expressed transcripts and proteins in different pairwise comparisons were subjected to overrepresentation analysis (ORA) using DAVID with gene ontology (GO) and KEGG pathway database to identify the enriched functional classes (Dennis et al., 2003). The enriched functional classes were then clustered in groups of terms and pathways with overlapping genes using the clustering function in DAVID.

5.4.3.1 Cellular responses at the transcriptome level

At the transcriptome level, enriched classes were clustered to groups related to transcription, MAPK signaling, and inflammation in 15 hpt vs. 0 hpt comparison (Table 5-2). At the late stage (48 hpt vs. 0 hpt), the number of differentially expressed transcripts increased and enriched functional classes clustered in groups involved in nucleosome assembly, host defense, and unfolded protein response (UPR). MAPK signaling related classes still formed a cluster while only an inflammatory response GO term was enriched

and listed in “unclustered” group. The enriched functional clusters in 48 hpt vs 0 hpt differed from Mock vs. 0 hpt, suggesting rAAV production elicited distinct transcriptome response compared to just plasmid transfection. However, cluster of host defense classes was also seen in mock transfection (Mock vs. 0 hpt) possibly reflecting general effects of DNA plasmid transfection.

Table 5-2. Clustering of functional classes identified in different paired comparisons of transcriptome by overrepresentation analysis using DAVID.ES: Enrichment score, GC: Gene count.

		Terms	ES	GC	%
15 hr vs 0 hr	Transcription	transcription, DNA-templated; regulation of transcription, DNA-templated	3.15	49	18.01
	MAPK signaling	MAP kinase tyrosine/serine/threonine phosphatase activity; MAPK signaling pathway; inactivation of MAPK activity	2.56	4	1.47
	Transcription	transcription from RNA polymerase II promoter; positive regulation of transcription from RNA polymerase II promoter; transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding	2.39	18	6.62
	Inflammation	response to muscle stretch; TNF signaling pathway; response to lipopolysaccharide; cellular response to calcium ion	2.07	4	1.47
	Unclustered	response to drug; response to xenobiotic stimulus, Human T-cell leukemia virus 1 infection; Herpes simplex virus 1 infection			
48 hr vs 0 hr	Nucleosome	Systemic lupus erythematosus; structural constituent of chromatin; Alcoholism; nuclear nucleosome; nucleosome; nucleosome assembly; Neutrophil extracellular trap formation; protein heterodimerization activity	4.87	42	2.64
	Host defense	Staphylococcus aureus infection; complement activation; Complement and coagulation cascades	3.66	15	0.88
	Unfolded protein response	positive regulation of transcription from RNA polymerase II promoter in response to endoplasmic reticulum stress; endoplasmic reticulum unfolded protein response; ATF6-mediated unfolded protein response	2.25	11	0.69
	MAPK signaling	protein tyrosine/threonine phosphatase activity; peptidyl-tyrosine dephosphorylation; negative regulation of MAPK cascade; MAP kinase tyrosine/serine/threonine phosphatase activity; inactivation of MAPK activity	1.80	9	0.57
	Unclustered	integrin-mediated signaling pathway; cell adhesion; inflammatory response; positive regulation of interleukin-6 production; Glycerin, serine and threonine metabolism			
Mock 48 hr vs 0 hr	ECM	extracellular matrix structural constituent; basement membrane; ECM-receptor interaction	4.03	26	1.75
	Host defense	Staphylococcus aureus infection; Complement and coagulation cascades; complement activation; complement activation, classical pathway; positive regulation of apoptotic cell clearance	2.36	12	0.81
	Cell motility	axonemal dynein complex; ATP-dependent microtubule motor activity, minus-end-directed; dynein light intermediate chain binding; dynein complex; cilium-dependent cell motility	2.26	8	0.54
	ECM dynamics	extracellular matrix structural constituent conferring tensile strength; collagen trimer; Protein digestion and absorption	2.21	14	0.94
	Unclustered	collagen fibril organization; inflammatory response; peptidyl-tyrosine dephosphorylation; response to calcium ion			

To assess viral components elicited responses, differentially expressed transcripts between 48 hpt and Mock samples were also subjected to ORA and functional class clustering. The enriched GO terms and KEGG pathways were clustered into groups related to nucleosome assembly (Cluster I), defense and immune response (Cluster II), UPR (Cluster III), and several unclustered classes related to growth regulation including p53 signaling and negative regulation of MAPK cascade were enriched (Table 5-3).

To further illustrate the similar but differential response in 48 hpt and in Mock, fold-changes of DEGs in 48 hpt vs. Mock comparison involved in functional classes identified by ORA are shown in Figure 5-9. As can be seen, more differentially expressed transcripts were in first and third quadrants (i.e., upregulated or downregulated in the same direction) than those in second and fourth quadrants (in opposite direction); but had different magnitude of response, especially those belonged to nucleosome assembly (Cluster I) and defense and immune response (Cluster II). On the other hand, DEGs in second and fourth quadrants were mostly involved in UPR (e.g., DERL3 and heat shock proteins), p53 signaling (e.g., cyclins, TP73, and DDB2), and negative regulation of MAPK cascade (e.g., DUSP genes). The data show that while plasmid transfection alone (Mock) and rAAV replication both elicited transcriptional changes in several functional classes, rAAV production did cause more prominent responses in nucleosome assembly and defense and immune response and specific responses in UPR, p53 signaling, and MAPK regulation.

Table 5-3. Functional class overrepresentation analysis (ORA) using DAVID and gene set enrichment analysis (GSEA) of 48 hpt vs. Mock comparison. Representative GO terms or KEGG pathways enriched at transcriptome and proteome levels.

48 hpt vs. Mock		Transcriptome		Proteome	
Cluster	GO Term and KEGG Pathway	ORA	GSEA (NES)	ORA	GSEA (NES)
I	Nucleosome Assembly	+			
I	Systemic Lupus Erythematosus		1.52		
II ^a	TNF Signaling Pathway			+	
II ^a	Interleukin 6 Production		1.90		1.28
II ^a	Cytokine Production Involved in Immune Response				1.77
II ^a	Chemokine Signaling Pathway				1.55
II ^b	Defense Response to Gram-Positive Bacterium	+			
II ^b	Complement Activation	+	1.85		
II ^b	Humoral Immune Response		1.59		1.78
II ^b	Regulation of Innate Immune Response				1.56
II ^b	Toll-Like Receptor Signaling Pathway			+	
II ^b	NOD-Like Receptor Signaling Pathway		1.25	+	1.62
II ^b	RIG-I Like Receptor Signaling Pathway				1.78
II ^b	Cytosolic DNA Sensing Pathway				1.41
III	Endoplasmic Reticulum Unfolded Protein Response	+			
III	Response to Unfolded Protein	+			
III	Protein Refolding		1.89		
U ^c	p53 Signaling Pathway	+			-1.13
U ^c	Cell Cycle				-1.12
U ^c	DNA Recombination			+	
U ^c	Nucleotide Excision Repair		-1.61		-1.61
U ^c	Postreplication Repair				-1.54
U ^d	Negative Regulation of MAPK Cascade	+			
U ^d	p38 MAPK cascade			+	
U ^d	MAPK Signaling Pathway		1.37		1.26
U ^d	Positive Regulation of ERK1 And ERK2 Cascade		1.82		1.34
U ^e	Endosomal Transport			+	-1.19
U ^e	Lysosome		-1.25		-1.34
U ^e	Endoplasmic Reticulum to Golgi Vesicle Mediated Transport		-1.60		
U ^e	Peroxisome		-1.42		

Footnote: +: Enriched in ORA ($p < 0.05$). NES: normalized enrichment score determined using GSEA. Cluster I: Nucleosome assembly. Cluster II: Defense and immune response. Cluster III: Unfolded protein response. U: Unclustered functional classes. a: inflammation. b: immune response. c: p53 signaling and DNA repair. d: MAPK signaling. e: protein processing.

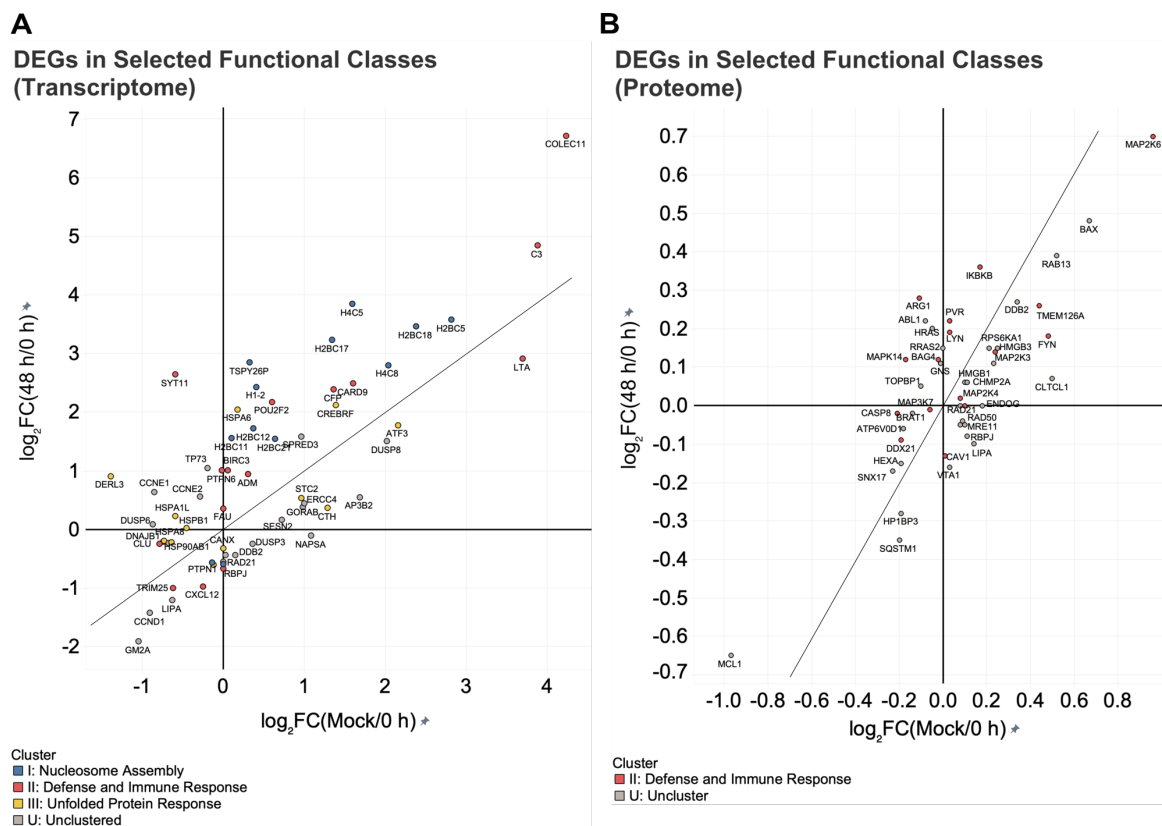


Figure 5-9. Differentially expressed genes involved in identified functional classes and their fold-changes in log₂ scale(A) Differentially expressed transcripts and (B) differentially expressed proteins with log₂ fold-changes of 48 hpt vs. 0 hpt and Mock vs. 0 hpt.

5.4.3.2 Cellular responses at the proteome level

While its gene coverage was less deep and dynamic range also not as wide, TMT proteomic analysis complemented transcriptomic analysis and provided further information on cellular responses at the protein level. ORA of differentially expressed proteins identified several classes related to protein translation and GTPase signaling as enriched throughout different stages of rAAV production (Table 5-4). Like in transcriptome, these functional classes were also enriched in Mock vs. 0 hpt. ORA was then performed on 48 hpt to Mock comparison. The enriched terms included several related to inflammatory and immune response, including TNF, Toll-like receptor (TLR) and NLR

signaling pathways (Cluster II), as well as unclustered others related to protein processing (e.g., endosomal transport), MAPK cascade, and DNA recombination (Table 5-3).

Table 5-4. Clustering of functional classes identified in different paired comparisons of proteome by overrepresentation analysis using DAVID.ES: Enrichment score, GC: Gene count.

	Terms		ES	GC	%
24 hr vs 0 hr	GTPase	small GTPase mediated signal transduction; GTPase activity; GTP binding; intracellular	4.22	26	5.23
	Muscle contraction	muscle filament sliding; muscle myosin complex; structural constituent of muscle; muscle contraction; myosin filament; microfilament motor activity; sarcomere; calmodulin binding; myosin complex	3.61	13	2.62
	Unclustered	retrograde vesicle-mediated transport, Golgi to ER; protein transport; Protein porcessing in endoplasmic reticulum; Protein export			
48 hr vs 0 hr	Translation	SRP-dependent cotranslational protein targeting to membrane; cytosolic large ribosomal subunit; viral transcription; rRNA processing; nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; Ribosome; translation; translational initiation; structural constituent of ribosome; ribosome	4.47	26	4.88
	GTPase	small GTPase mediated signal transduction; GTPase activity; GDP binding; GTP binding; intracellular	4.34	27	5.07
	Muscle contraction	muscle filament sliding; structural constituent of muscle; muscle myosin complex; muscle contraction	3.87	11	2.06
	Unclustered	cell division; defense response to fungus; positive regulation of ubiquitin protein ligase activity; retrograde vesicle-mediated transport, Golgi to ER			
72 hr vs 0 hr	Translation	SRP-dependent cotranslational protein targeting to membrane; Ribosome; nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; viral transcription; translational initiation; cytosolic large ribosomal subunit; rRNA processing; structural constituent of ribosome; translation; ribosome; cytosolic small ribosomal subunit	9.53	37	6.62
	GTPase	GTPase activity; small GTPase mediated signal transduction; GTP binding	4.57	30	5.37
	Kinetochores	sister chromatid cohesion; kinetochore; condensed chromosome kinetochore	2.11	18	3.22
	Unclustered	p53 signaling; Cell cycle; AMPK signaling pathway			
Mock vs 0 hr	Translation	cytosolic large ribosomal subunit; SRP-dependent cotranslational protein targeting to membrane; viral transcription; nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; translation; structural constituent of ribosome; rRNA processing; Ribosome; translational initiation; ribosome	5.13	22	4.13
	GTPase	small GTPase mediated signal transduction; GTPase activity; GDP binding; GTP binding; intracellular	3.46	25	4.69
	Muscle contraction	muscle filament sliding; muscle myosin complex; muscle contraction; microfilament motor activity; myosin filament; structural constituent of muscle; sarcomere	3.16	11	2.06
	Unclustered	cell division; positive regulation of ubiquitin protein ligase activity; protein transport;			

GSEA was also performed on proteome of 48 hpt vs. Mock (Table 5-3). Like in ORA of proteome, several enriched functional classes were related to immune and host

response, including humoral immune response, cytokine production, innate immune response, RIG-I like receptor (RLR), NLR signaling pathways, and cytosolic DNA sensing pathway (Cluster II). Unclustered p53 signaling, DNA repair, and MAPK signaling were also enriched. Differentially expressed proteins involved in functional classes identified by either ORA or GSEA were shown in Figure 5-9B.

5.4.4 Cellular Responses – Integrated View

Integrating functional analysis at both transcriptome and proteome levels, one saw a general trend of enrichment of functions related to host defense and immune response (Cluster II) (Table 5-3). The twelve functional classes in the defense and immune responses cluster fell into two groups: inflammation (II^a) and immune response (II^b). A key response by host cell to virus infection is inflammatory response by cytokine secretion to communicate to other cells about the virus infection and to elicit antiviral responses in other cells (Mogensen & Paludan, 2001). Although the triple transfection is somewhat different from virus infection, the replication of viral genome and the expression of AAV and Ad viral proteins also triggered inflammatory reactions at transcript and protein levels.

5.4.4.1 Host immune response - complement activation and pattern recognition receptors (PRRs)

Virus replication elicited host cell immune responses (Cluster II^b) includes reactions related to complement activation and pattern recognition receptor (PRR) family including signaling pathways of Toll-like receptor (TLR), NOD-like receptor (NLR) and RIG-I like receptor (RLR) (Ma et al., 2018). C3 (Complement component 3) and CFP (Complement factor P) were upregulated at the transcript level in 48 hpt/Mock comparison (Figure 5-9A). C3 was reported to associate with AAV2 capsids in AAV-transduced cells

(Zaiss et al., 2008). Its role in AAV-activated humoral immune responses was further shown by the reported weaker humoral immunity against AAV2 in C3-deficient mice. CFP was reported to be a positive regulator of alternative complement pathway (Narni-Mancinelli et al., 2017). It is plausible that complement activation is part of cellular responses to clear AAV and temper AAV replication.

Membrane-associated receptors in PRR are responsible for detecting pathogen-associated molecular patterns (PAMPs) motifs like viral genome, RNA and proteins and initiating a series of innate immune responses (Ma et al., 2018). AAV transduction was reported to activate RLR signaling and upregulation of the dsRNA sensor MDA5 (Shao et al., 2018). More specifically, capsids of AAV, as PAMPs, was reported to cause the expression of proinflammatory cytokines like TNF- α , IL-6, and IL-8 through TLR signaling (Hösel et al., 2012), and AAV DNA was shown to be recognized by TLR9 triggering the TLR-MyD88 signaling pathway and type I interferon production (Muhuri et al., 2021; Zhu et al., 2009).

Triple transfection also introduced Ad helper genes to cells. Ad non-coding VA RNA could be recognized by RLR that activated IRF3 followed by type I interferon production (Minamitani et al., 2011). A transcriptomic analysis of triple transfection also identified genes involved in defense and immune responses, such as C3, FOS, IL-17, that were differentially expressed. The expression of type I interferon and multiple interferon-stimulated genes, including IFIT2 and OASL, also increased during rAAV production (Chung et al., 2023).

5.4.4.2 *Unfolded protein response*

With the very high-level expression of viral and GOI proteins, it was not surprised to see transcripts of genes involved in protein folding and UPR were differentially expressed. Among transcripts in this cluster (Figure 5-9A), DERL3 was involved in the degradation of misfolded glycoproteins in the ER and positively correlated with UPR (Oda et al., 2006). Many differentially expressed transcripts were members in heat shock protein family (e.g., HSPA6, HSPA8, HSPB1, DNAJB1), the molecular chaperons responsible for refolding misfolded proteins. CREBRF is a CREB3 regulatory factor and negative regulator of UPR. It has been reported that UPR reduced global protein synthesis and induced apoptosis (Prasad & Greber, 2021). Through its global effects on protein synthesis and inducing AAV capsid protein degradation (Balakrishnan et al., 2013), UPR can lead to lower rAAV productivity.

5.4.4.3 *MAPK cascade*

MAPK cascade and negative regulation of MAPK cascade were among unclustered functional classes. DUSP (Dual-specificity phosphatase) genes (Figure 5-9A) which dephosphorylate MAPKs to regulate the duration or magnitude of MAPK activity were among those differentially expressed transcripts. Many genes differentially expressed at the proteome level in MAPK signaling, such as MAP2K3, MAP2K6 and MAPK14 (Figure 5-9B), were involved in the p38-MAPK pathway. MAPK14 serves as an integration point of a number of responses such as inflammation and activation of transcription factors (Zarubin & Han, 2005). Their activation potentially promotes virus replication by promoting viral protein synthesis, viral genome replication, and suppressing immune response and apoptosis (Kumar et al., 2018). Inhibiting p38-MAPK signaling by

using SB203580 inhibited replication of herpes simplex virus and encephalomyocarditis virus (Kumar et al., 2018).

5.4.4.4 Cell cycle regulation

Many unclustered enriched functional classes were related to growth control (U^c), including p53 signaling and cell cycle (Table 5-3). Viruses use cellular machinery to replicate their genome and reproduce. Upon infection, they interfere with cell's cell cycle control to keep it at a stage that favors viral genome replication (Bagga & Bouchard, 2014). Many genes in the cyclin family, such as CCND1, CCNE1, and CCNE2, were differentially expressed (Figure 5-9A). Through cyclin-dependent kinases (CDKs), cyclins drive the cell cycle progression. Rep78 protein of AAV has been reported to alter CDK activity, increase the p21 (CDK inhibitor) level, and inhibit CDC25A that caused cell cycle arrest in S phase (Berthet et al., 2005; Saudan et al., 2000). Arresting in S/G2 phase has been reported to promote AAV DNA replication (Franzoso et al., 2017). A recent report showed that addition of nocodazole, an anti-mitotic agent, increased rAAV production by triple transfection, highlighting intervening cell cycle regulation to boost AAV genome replication (Scarrott et al., 2023).

5.4.4.5 DNA damage and repair response

Some components of AAV can cause a DNA damage response in infected cells. Rep78 protein has endonuclease activity, which can nick cellular chromatin, generate double-strand DNA breaks, and elicit a DNA damage response mediated by ATM (ataxia telangiectasia mutated), and triggers S phase cell cycle arrest (Berthet et al., 2005). Additionally, AAV ITR with hairpin structure could also elicit a DNA damage response (Raj et al., 2001). ATM interacts with MRN complex (MRE11, RAD50, and NBS1) to

activate an ATM signaling-mediated DNA damage response. DNA viruses replicate their genome inside the nucleus in localized replication compartment (Knipe et al., 2022). Activation of ATM signaling leads to recruiting of cellular DNA replication machinery, including DNA polymerase and DNA damage-signaling proteins, such as TOPBP1 and BRCA1 proteins, to form replication compartment which are required for AAV replication (Jurvansuu et al., 2005; Weitzman & Fradet-Turcotte, 2018). In addition to playing an important role in DNA damage responses and DNA repair, MRN complex also plays a role in restraining AAV replication as part of host cell defense. It has nuclease activity and can degrade AAV genome. Two adenovirus helper proteins, E1B55K and E4orf6, form a complex, which induces degradation of MRN to suppress MRN's nuclease activity, and increase AAV production (Schwartz et al., 2007). Without the presence of Ad, MRN complex accumulates at AAV replication foci (Collaco et al., 2009). Several components in the pathway, including DDB2, TOPBP1, and BRAT1, were differentially expressed at the protein level (Figure 5-9B). Two components of MRN complex, MRE11 and RAD50, were downregulated at the protein level.

5.4.5 Modulation of Host Responses for Enhancing rAAV Production

With enriched functional clusters identified, one is tempted to explore the possibility of genetic or chemical intervention of the involved pathways to increase rAAV productivity. We tested eight small molecules and four PRR nucleic acid agonists to assess their impact on rAAV yield by treating HEK293 cells with these agents followed by triple transfection (Table 5-5).

Table 5-5. Small molecules, and agonists of pathogen pattern recognition signaling used to evaluate the effect of intervention of pathways of interest on rAAV production.

Small Molecules			
Target	Name	Cat. No	Mode of Action
RIG-I/Toll-Like Receptor Signaling Pathway	MRT 67307	SML0702-5MG	A potent, reversible kinase inhibitor specifically blocking the function of TBK1 and IKK ϵ , thereby preventing the phosphorylation of IRF3 and subsequent expression of ISGs.
PKC and NF- κ B	PMA	tlrl-pma	A specific activator of Protein Kinase C (PKC) which involves a wide range of signaling pathways including cell proliferation and activation of NF- κ B.
NF- κ B	QNZ	481407-1MG	Shows potent inhibitory activity toward both NF- κ B activation and TNF- α production. The mechanism is unknown.
p38 MAPK	SB203580	559389-1MG	Inhibits p38 MAPK catalytic activity by binding to the ATP binding pocket.
Apoptosis	Z-IETD-FMK	218759-250UG	The tetrapeptide IETD binds preferentially to the active site of CASP-8. (Caspase-8 inhibitor)
	V5	196810-5MG	Interact with Bax, prevent its conformational change and mitochondrial translocation. Acts as an effective Inhibitor for Bax-mediated apoptosis.
DNA Damage Response	KU-55933	SML1109-5MG	Inhibits ATM kinase by binding to its catalytic site that prevents ATM-dependent phosphorylation event and interferes DNA damage signaling.
	Mirin	M9948-5MG	Disrupts nuclease activity of MRE11 and thus, the ability to repair DNA double strand breaks. Inhibits MRN complex-dependent ATM signaling.
Nucleic Acid Agonists			
Target	Name	Cat. No	Mode of Action
Pattern Recognition Receptor (PRR)	3p-hpRNA	tlrl-hprna	A 5' triphosphate hairpin RNA self-anneals to form secondary structures such as hairpin or panhandle conformations. These structural features, which distinguish viral RNA from mammalian RNA, are recognized by RIG-I which triggers antiviral immune response.
	Poly(dA:dT), dsDNA	tlrl-patn	A repetitive synthetic double-stranded DNA sequence of poly(dA-dT):poly(dT-dA) recognized by cytosolic DNA sensors triggers the production of interferons.
	Poly(I:C) (HMW)	tlrl-pic	A synthetic analog of dsRNA recognized by RIG-I/MDA5 and TLR3 triggers the production of interferons, AP-1 and NF- κ B.
	Poly(I:C) (LMW)	tlrl-picw	A synthetic analog of dsRNA recognized by RIG-I/MDA5 and TLR3 triggers the production of interferons, AP-1 and NF- κ B.

An inhibitor of PRR signaling pathway, MRT67307, boosted rAAV productivity (Figure 5-10A). MRT 67307 blocks the function of TBK1 (TANK-binding kinase 1)/IKK ϵ (I-kappa-B kinase epsilon) complex, which is a hub in RLR and TLR signaling pathways (Figure 5-11A), thus prevents the phosphorylation and activation of IRF3 (interferon regulatory factor 3) and the following expression of ISGs (interferon stimulated genes) (Clark et al., 2011; Fitzgerald et al., 2003). Four nucleic acid agonists of PRRs greatly decreased rAAV titers (Figure 5-10B), suggesting that attenuating host cell antiviral response can be a means of increasing rAAV productivity. Another molecule which also had positive effect on rAAV titer was an activator of Protein kinase C (PKC) called PMA (Phorbol 12-myristate 13-acetate), which activates the transcription factor NF κ B (Figure

5-11B) (Hellweg et al., 2006). When an inhibitor of NF κ B transcriptional activation, QNZ, was used, rAAV titer decreased. Intervention of other functional pathways including p38 MAPK, apoptosis, and DNA damage response did not have any positive effect (Figure 5-10A).

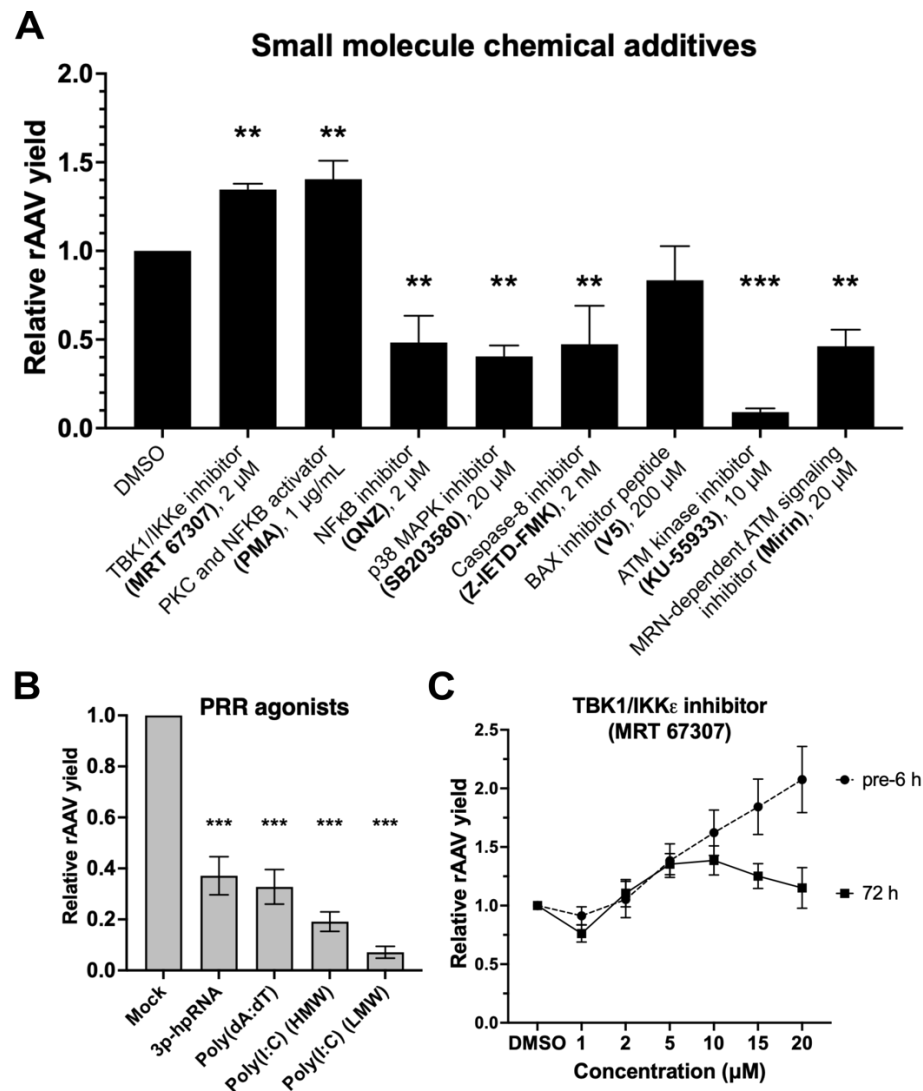


Figure 5-10. Modulation of host responses for enhancing rAAV production.(A) The effect of small molecule chemical additives and (B) pattern recognition receptor agonists on rAAV yield. (C) Dose-dependent effect of MRT67307 on rAAV production. For A and B, data represent mean and standard deviation of three independent replicates determined by qPCR. For C, data represent mean and standard deviation of transducing units calculated from GFP-positive cell counts from multiple images. For A, small molecules were added at 0 h and kept in culture throughout the 72 h-period, and rAAV yield was in reference to

mock control with 1% DMSO. For B, cells were transfected with PRR agonists followed by triple transfection, and rAAV yield was in reference to mock transfection. For C, MRT67307 was added 6 h before transfection and removed at 0 h (pre-6 h) or used as described in A (72 h). *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ as determined by a one-way analysis of variance (ANOVA) multiple-comparisons test.

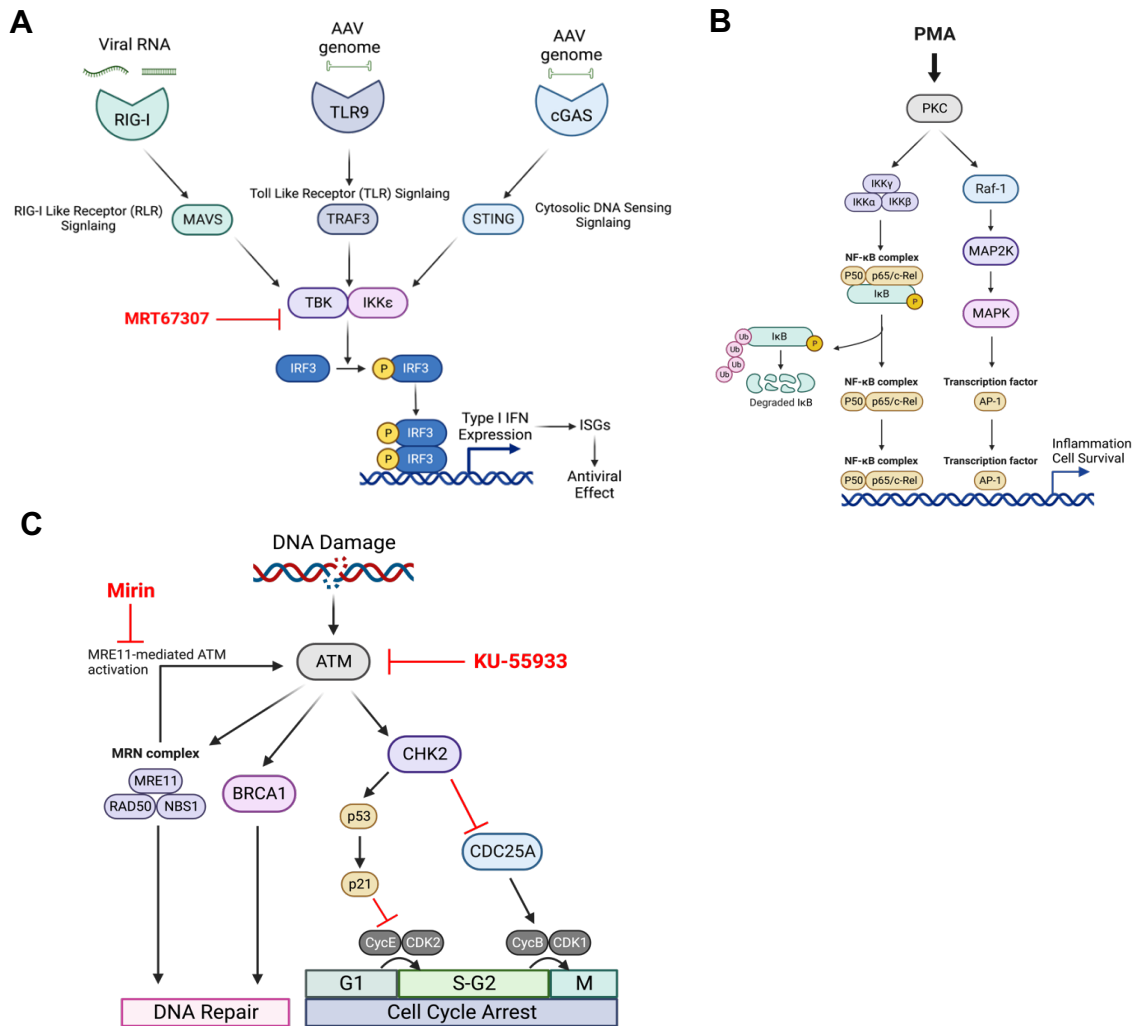


Figure 5-11. Illustration of modulation of signaling pathways using small molecules.(A) Pathogen recognition receptor signaling and inhibition of TBK1/IKKε by MRT67307. (B) Activation of protein kinase C and NF-κB signaling by PMA. (C) Inhibition of ATM kinase and MRE11-mediated ATM signaling by KU-55933 and Mirin respectively.

ATM kinase is activated by DNA double-strand break and leads to arrest at DNA damage checkpoint (Figure 5-11C). The inhibition of ATM kinase and MRE11-dependent ATM signaling, using KU55933 and Mirin, respectively, decreased rAAV production

(Figure 5-10A). This finding is consistent with earlier reports involving the use of KU55933 or knockdown of ATM (Collaco et al., 2009; Jurvansuu et al., 2005).

To seek for a higher titer and further validate our observation, dose-dependent responses of chemicals with positive effect, MRT67307 and PMA, were further determined (Figure 5-10C and Figure 5-12A). In addition to originally used incubation condition (added at 0 h and kept all the way through), a shorter exposure time (cells were treated with chemicals for 6 h followed by medium replacement and triple transfection) was also tested. With the new incubation condition, titers increased along with higher doses of MRT67307, and up to 2-fold increase in rAAV titer was seen while no notable dose-dependent response and difference between two culture conditions were seen for PMA. On the other hand, higher doses of NF κ B and ATM inhibitors resulted in stronger negative effect on rAAV yield (Figure 5-12B, C).

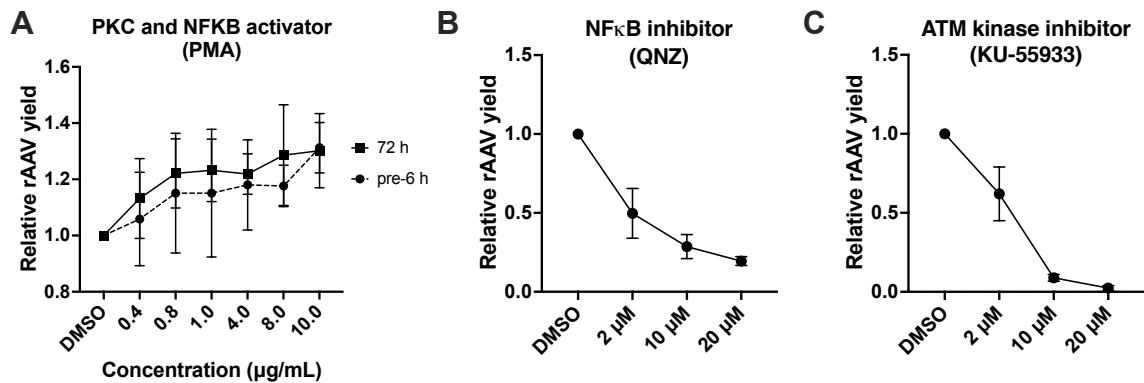


Figure 5-12. Dose-dependent effect of (A) PMA, (B) NF- κ B inhibitor, and (C) ATM inhibitor on rAAV production determined by transduction assay. Data represent mean transducing units calculated from GFP-positive cell counts from multiple images. rAAV yield was in reference to Mock control with 1% DMSO. Small molecules were added at 0 h and kept in culture throughout the 72 h-period. For A, PMA was added 6 h before transfection and removed at 0 h (pre-6 h) or used as previously described (72 h).

A challenge of such intervention is that the activity of a functional feature may be promoting as well as suppressing virus replication. As ATM-mediated DNA damage response is essential for AAV replication, it can also cause the degradation of AAV genome by MRN complex to decrease AAV titer, unless it is balanced by helper protein E1B55K/E4orf6 induced degradation of MRN complex. Hence the effect of intervening DNA damage pathway could be context dependent.

5.5 Concluding Remarks

We have integrated transcriptomic and proteomic analysis to gain fresh insight into a commonly used transient transfection process of rAAV production. The quantitative and kinetic characterization of viral and host cell proteins suggests that packaging of rAAV genome into assembled capsid is inefficient. Integrative analysis of host transcriptome and proteome indicated defense and immune response, UPR, and growth regulation during rAAV production were altered after triple transfection. The multi-omics analysis-guided testing of small molecules demonstrated the potential of enhancing rAAV productivity by modulation of host cell responses. The new understanding of rAAV production kinetics and host responses gained in this study can potentially guide further enhancement of rAAV productivity in HEK293 cells.

6 Summary and Concluding Remarks

Recombinant adeno-associated virus (rAAV)-based gene therapy has emerged as one of the important therapeutic tools to treat genetic diseases, as demonstrated by several clinical successes. However, the increasing demand for rAAV in late-phase studies and clinical applications pose challenges to large-scale manufacturing. Additionally, the quality of vectors, specifically full particle content, produced by current production processes needs improvement to alleviate stress in downstream purification processes. Complex rAAV production involves viral protein synthesis, genome replication, capsid assembly, and genome encapsidation, relying on both viral components and host cell machinery. This complexity complicates the development and improvement of rAAV production. To overcome these challenges, we innovated an inducible rAAV producer cell line as a novel production platform. We have gained insights into rAAV production through comprehensive examinations of the expression dynamics of viral genes and the host cellular responses to viral replication. Both the rAAV yield and quality produced using inducible cell lines were greatly enhanced through design-build-characterize-learn cycles.

This thesis presents the development of inducible rAAV producer cell lines through a synthetic biology approach. Genetic modules containing rAAV genome with ITR-flanked GOI, inducible *rep*, *cap* and helper genes were integrated into the HEK293 cell genome via a transposon system. The synthetic cell lines were capable of producing rAAV upon small molecule induction. Moreover, with independent control over Rep and Cap protein expression, full particle content could be boosted to a high level.

To further enhance rAAV productivity in synthetic cell lines, we adopted multiple physical assays and omics technologies to identify bottlenecks in rAAV production and study the kinetics of viral components. After identifying wasteful GOI expression as a limiting factor, we replaced constitutive promoter with a LacSwitch inducible promoter to repress GOI expression. This new design led to increases in the number of integrated copies of rAAV genome, the VG titer, and full particle content. In addition, based on the time-course protein dynamics measured by targeted proteomics, we discovered that proteasomal degradation of Cap proteins might limit rAAV production. Cell lines treated with proteasome inhibitors had substantially higher titers.

By comparing abundance of viral components, we found out that synthetic cell lines could amplify virus genomes to a very high level but had low capsid titers, which might limit virus packaging. We also confirmed that *cap* gene expression was lower at both the transcript and protein levels, when compared to the triple transfection system. Based on these findings, we hypothesized that *cap* gene expression was one of the limiting factors. To address this limitation, we introduced more copies of the *cap* gene into previously developed cell lines through sequential transfection, successfully isolating clones capable of producing 10^5 vector genome per cell. Manipulation of inducer concentrations or induction time profiles improved full particle content without a loss in full particle titers. The integrated view of kinetic profiles of viral components suggested that the slow genome amplification and fast capsid assembly might have an impact on product quality. Finally, a high producer cell line was adapted to suspension culture to demonstrate its scalability.

In addition to studying viral genes, we conducted a comprehensive multi-omics analysis of cellular physiological changes during transfection-based rAAV production at

the transcriptome and proteome levels. We observed upregulation of inflammatory responses, defense and immune responses, and unfolded protein responses. In addition, altered cell growth regulations, including MAPK signaling, cell cycle, p53 signaling, and DNA damage and repair, might play a critical role in viral replication. Guided by functional class analysis, we explored the modulation of signaling pathways, such as pathogen recognition receptor signaling pathways, using small molecule chemicals. This approach successfully enhanced transfection-based rAAV production. Our multi-omics analysis provides valuable guidance for further improvement of rAAV production.

There are numerous future directions we can pursue to advance our technology. We observed that amplified virus genome copies became limited in the most recent cell lines, once capsid titers were highly increased. Since we only used a minimal set of viral genes for synthetic cell lines, we can consider including additional viral genes that could benefit rAAV replication, such as the E2F activator E4orf6/7 from adenovirus or the UL30/UL42 DNA polymerase complex from HSV. Incorporating such genes can potentially push the titers and full particle content higher. On the other hand, as a factory, host cells play a vital role in viral vector production. Host cells have developed strategies to fight against pathogens. Thus, to block defense mechanisms and to create a favorable environment for viral replication, host cell genome engineering via CRISPR/Cas9 should be considered to attenuate antiviral responses and apoptosis and to enhance DNA replication capacity and energy metabolism. Finally, when applying cell lines in industrial processes, we can optimize cell culture conditions in upstream processes using various strategies. Better media design and feeding strategies guided by empirical and/or mechanistic models based on quantitative metabolite measurements can potentially lead to higher productivity.

Perfusion culture, which allows for high-cell-density culture by replenishing nutrients and removing inhibitory compounds, could elevate titer further. A bright promise of rAAV production using synthetic rAAV producer cell lines can be seen with a variety of available tools and upcoming new technology.

7 Bibliography

- Alazard-Dany, N., Nicolas, A., Ploquin, A., Strasser, R., Greco, A., Epstein, A. L., Fraefel, C., & Salvetti, A. (2009). Definition of Herpes Simplex Virus Type 1 Helper Activities for Adeno-Associated Virus Early Replication Events. *PLOS Pathogens*, 5(3), e1000340. <https://doi.org/10.1371/journal.ppat.1000340>
- Allen, J. M., Debelak, D. J., Reynolds, T. C., & Miller, A. D. (1997). Identification and elimination of replication-competent adeno-associated virus (AAV) that can arise by nonhomologous recombination during AAV vector production. *Journal of virology*, 71(9), 6816-6822. <https://doi.org/10.1128/jvi.71.9.6816-6822.1997>
- Anguela, X. M., & High, K. A. (2019). Entering the Modern Era of Gene Therapy. *Annual Review of Medicine*, 70(1), 273-288. <https://doi.org/10.1146/annurev-med-012017-043332>
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1), 25-29. <https://doi.org/10.1038/75556>
- Au, H. K. E., Isalan, M., & Mielcarek, M. (2022). Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings. *Frontiers in Medicine*, 8. <https://doi.org/10.3389/fmed.2021.809118>
- Ayuso, E., Mingozzi, F., & Bosch, F. (2010). Production, Purification and Characterization of Adeno-Associated Vectors. *Current Gene Therapy*, 10(6), 423-436. <https://doi.org/10.2174/156652310793797685>
- Ayuso, E., Mingozzi, F., Montane, J., Leon, X., Anguela, X. M., Haurigot, V., Edmonson, S. A., Africa, L., Zhou, S., High, K. A., Bosch, F., & Wright, J. F. (2010). High AAV vector purity results in serotype- and tissue-independent enhancement of transduction efficiency. *Gene Therapy*, 17(4), 503-510. <https://doi.org/10.1038/gt.2009.157>
- Bachhav, B., de Rossi, J., Llanos, C. D., & Segatori, L. (2023). Cell factory engineering: Challenges and opportunities for synthetic biology applications. *Biotechnology and Bioengineering*, 120(9), 2441-2459. <https://doi.org/10.1002/bit.28365>
- Bagga, S., & Bouchard, M. J. (2014). Cell Cycle Regulation During Viral Infection. In E. Noguchi & M. C. Gadaleta (Eds.), *Cell Cycle Control: Mechanisms and Protocols*

- (pp. 165-227). Springer New York. https://doi.org/10.1007/978-1-4939-0888-2_10
- Balakrishnan, B., Sen, D., Hareendran, S., Roshini, V., David, S., Srivastava, A., & Jayandharan, G. R. (2013). Activation of the Cellular Unfolded Protein Response by Recombinant Adeno-Associated Virus Vectors. *PLoS One*, 8(1), e53845. <https://doi.org/10.1371/journal.pone.0053845>
- Balasubramanian, S., Wurm, F. M., & Hacker, D. L. (2016). Multigene expression in stable CHO cell pools generated with the piggyBac transposon system. *Biotechnology Progress*, 32(5), 1308-1317. <https://doi.org/10.1002/btpr.2319>
- Banaszynski, L. A., Chen, L.-C., Maynard-Smith, L. A., Ooi, A. G. L., & Wandless, T. J. (2006). A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell*, 126(5), 995-1004. <https://doi.org/10.1016/j.cell.2006.07.025>
- Berthet, C., Raj, K., Saudan, P., & Beard, P. (2005). How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proceedings of the National Academy of Sciences*, 102(38), 13634-13639. <https://doi.org/10.1073/pnas.0504583102>
- Brough, D. E., Lizonova, A., Hsu, C., Kulesa, V. A., & Kovesdi, I. (1996). A gene transfer vector-cell line system for complete functional complementation of adenovirus early regions E1 and E4. *Journal of virology*, 70(9), 6497-6501. <https://doi.org/10.1128/jvi.70.9.6497-6501.1996>
- Chadeuf, G., Favre, D., Tessier, J., Provost, N., Nony, P., Kleinschmidt, J., Moullier, P., & Salvetti, A. (2000). 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. *The Journal of Gene Medicine*, 2(4), 260-268. [https://doi.org/10.1002/1521-2254\(200007/08\)2:4<260::AID-JGM111>3.0.CO;2-8](https://doi.org/10.1002/1521-2254(200007/08)2:4<260::AID-JGM111>3.0.CO;2-8)
- Chejanovsky, N., & Carter, B. J. (1989). Mutagenesis of an AUG codon in the adeno-associated virus rep gene: Effects on viral DNA replication. *Virology*, 173(1), 120-128. [https://doi.org/10.1016/0042-6822\(89\)90227-4](https://doi.org/10.1016/0042-6822(89)90227-4)
- Chen, C.-L., Clark, R., Chen, R., Schnepf, B. C., & Johnson, P. R. (2002). Generation of Stable 293 Cell Lines for the Production of rAAV Vectors. *Molecular Therapy*, 5(5), S47. [https://doi.org/10.1016/S1525-0016\(16\)42967-9](https://doi.org/10.1016/S1525-0016(16)42967-9)
- Chen, S.-J., Johnston, J., Sandhu, A., Bish, L. T., Hovhannisyan, R., Jno-Charles, O., Sweeney, H. L., & Wilson, J. M. (2013). Enhancing the Utility of Adeno-Associated Virus Gene Transfer through Inducible Tissue-Specific Expression. *Human Gene Therapy Methods*, 24(4), 270-278. <https://doi.org/10.1089/hgtb.2012.129>

- Chung, C.-H., Murphy, C. M., Wingate, V. P., Pavlicek, J. W., Nakashima, R., Wei, W., McCarty, D., Rabinowitz, J., & Barton, E. (2023). Production of rAAV by plasmid transfection induces antiviral and inflammatory responses in suspension HEK293 cells. *Molecular Therapy - Methods & Clinical Development*, 28, 272-283. <https://doi.org/10.1016/j.omtm.2023.01.002>
- Clark, K., Pegg, M., Plater, L., Sorcek, Ronald J., Young, Erick R. R., Madwed, Jeffrey B., Hough, J., McIver, Edward G., & Cohen, P. (2011). Novel cross-talk within the IKK family controls innate immunity. *Biochemical Journal*, 434(1), 93-104. <https://doi.org/10.1042/BJ20101701>
- Clark, K. R., Voulgaropoulou, F., Fraley, D. M., & Johnson, P. R. (1995). Cell Lines for the Production of Recombinant Adeno-Associated Virus. *Human Gene Therapy*, 6(10), 1329-1341. <https://doi.org/10.1089/hum.1995.6.10-1329>
- Clément, N., & Grieger, J. C. (2016). Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Molecular Therapy - Methods & Clinical Development*, 3, 16002. <https://doi.org/10.1038/mtm.2016.2>
- Clément, N., Knop, D. R., & Byrne, B. J. (2009). Large-Scale Adeno-Associated Viral Vector Production Using a Herpesvirus-Based System Enables Manufacturing for Clinical Studies. *Human Gene Therapy*, 20(8), 796-806. <https://doi.org/10.1089/hum.2009.094>
- Collaco, R. F., Bevington, J. M., Bhargava, V., Kalman-Maltese, V., & Trempe, J. P. (2009). Adeno-associated virus and adenovirus coinfection induces a cellular DNA damage and repair response via redundant phosphatidylinositol 3-like kinase pathways. *Virology*, 392(1), 24-33. <https://doi.org/10.1016/j.virol.2009.06.005>
- Crystal, R. G. (2014). Adenovirus: the first effective in vivo gene delivery vector. *Human Gene Therapy*, 25(1), 3-11. <https://doi.org/10.1089/hum.2013.2527>
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, 4(9), R60. <https://doi.org/10.1186/gb-2003-4-9-r60>
- Destro, F., Joseph, J., Srinivasan, P., Kanter, J. M., Neufeld, C., Wolfrum, J. M., Barone, P. W., Springs, S. L., Sinskey, A. J., Cecchini, S., Kotin, R. M., & Braatz, R. D. (2023). Mechanistic modeling explains the production dynamics of recombinant adeno-associated virus with the baculovirus expression vector system. *Molecular Therapy - Methods & Clinical Development*, 30, 122-146. <https://doi.org/10.1016/j.omtm.2023.05.019>
- Dickerson, R., Argento, C., Pieracci, J., & Bakhshayeshi, M. (2021). Separating Empty

- and Full Recombinant Adeno-Associated Virus Particles Using Isocratic Anion Exchange Chromatography. *Biotechnology Journal*, 16(1), 2000015. <https://doi.org/10.1002/biot.202000015>
- Dobrowsky, T., Gianni, D., Pieracci, J., & Suh, J. (2021). AAV manufacturing for clinical use: Insights on current challenges from the upstream process perspective. *Current Opinion in Biomedical Engineering*, 20, 100353. <https://doi.org/10.1016/j.cobme.2021.100353>
- Doerr, A. (2012). Targeting with PRM. *Nature Methods*, 9(10), 950-950. <https://doi.org/10.1038/nmeth.2193>
- Dunbar, C. E., High, K. A., Joung, J. K., Kohn, D. B., Ozawa, K., & Sadelain, M. (2018). Gene therapy comes of age. *Science*, 359(6372), eaan4672. <https://doi.org/10.1126/science.aan4672>
- El Andari, J., & Grimm, D. (2021). Production, Processing, and Characterization of Synthetic AAV Gene Therapy Vectors. *Biotechnology Journal*, 16(1), 2000025. <https://doi.org/10.1002/biot.202000025>
- El Karoui, M., Hoyos-Flight, M., & Fletcher, L. (2019). Future Trends in Synthetic Biology—A Report [Perspective]. *Frontiers in Bioengineering and Biotechnology*, 7. <https://doi.org/10.3389/fbioe.2019.00175>
- Elmore, Z. C., Patrick Havlik, L., Oh, D. K., Anderson, L., Daaboul, G., Devlin, G. W., Vincent, H. A., & Asokan, A. (2021). The membrane associated accessory protein is an adeno-associated viral egress factor. *Nature Communications*, 12(1), 6239. <https://doi.org/10.1038/s41467-021-26485-4>
- Emmerling, V. V., Pegel, A., Milian, E. G., Venereo-Sanchez, A., Kunz, M., Wegele, J., Kamen, A. A., Kochanek, S., & Hoerer, M. (2016). Rational plasmid design and bioprocess optimization to enhance recombinant adeno-associated virus (AAV) productivity in mammalian cells. *Biotechnology Journal*, 11(2), 290-297. <https://doi.org/10.1002/biot.201500176>
- Fan, P.-D., & Dong, J.-Y. (1997). Replication of rep-cap Genes Is Essential for the High-Efficiency Production of Recombinant AAV. *Human Gene Therapy*, 8(1), 87-98. <https://doi.org/10.1089/hum.1997.8.1-87>
- Farson, D., Harding, T. C., Tao, L., Liu, J., Powell, S., Vimal, V., Yendluri, S., Koprivnikar, K., Ho, K., Twitty, C., Husak, P., Lin, A., Snyder, R. O., & Donahue, B. A. (2004). Development and characterization of a cell line for large-scale, serum-free production of recombinant adeno-associated viral vectors. *The Journal of Gene Medicine*, 6(12), 1369-1381. <https://doi.org/10.1002/jgm.622>

- Fitzgerald, K. A., McWhirter, S. M., Faia, K. L., Rowe, D. C., Latz, E., Golenbock, D. T., Coyle, A. J., Liao, S.-M., & Maniatis, T. (2003). IKK ϵ and TBK1 are essential components of the IRF3 signaling pathway. *Nature Immunology*, 4(5), 491-496. <https://doi.org/10.1038/ni921>
- François, A., Bouzelha, M., Lecomte, E., Broucque, F., Penaud-Budloo, M., Adjali, O., Moullier, P., Blouin, V., & Ayuso, E. (2018). Accurate Titration of Infectious AAV Particles Requires Measurement of Biologically Active Vector Genomes and Suitable Controls. *Molecular Therapy - Methods & Clinical Development*, 10, 223-236. <https://doi.org/10.1016/j.omtm.2018.07.004>
- Franzoso, F. D., Seyffert, M., Vogel, R., Yakimovich, A., de Andrade Pereira, B., Meier, A. F., Sutter, S. O., Tobler, K., Vogt, B., Greber, U. F., Büning, H., Ackermann, M., & Fraefel, C. (2017). Cell Cycle-Dependent Expression of Adeno-Associated Virus 2 (AAV2) Rep in Coinfections with Herpes Simplex Virus 1 (HSV-1) Gives Rise to a Mosaic of Cells Replicating either AAV2 or HSV-1. *Journal of virology*, 91(15). <https://doi.org/10.1128/JVI.00357-17>
- Galibert, L., Jacob, A., Savy, A., Dickx, Y., Bonnin, D., Lecomte, C., Rivollet, L., Sanatine, P., Boutin Fontaine, M., Le Bec, C., & Merten, O.-W. (2021). Monobac System—A Single Baculovirus for the Production of rAAV. *Microorganisms*, 9(9). <https://doi.org/10.3390/microorganisms9091799>
- Gao, K., Li, M., Zhong, L., Su, Q., Li, J., Li, S., He, R., Zhang, Y., Hendricks, G., Wang, J., & Gao, G. (2014). Empty virions in AAV8 vector preparations reduce transduction efficiency and may cause total viral particle dose-limiting side effects. *Molecular Therapy - Methods & Clinical Development*, 1. <https://doi.org/10.1038/mtm.2013.9>
- Gonçalves, M. A. F. V. (2005). Adeno-associated virus: from defective virus to effective vector. *Virology Journal*, 2(1), 43. <https://doi.org/10.1186/1743-422X-2-43>
- Grieger, J. C., Soltys, S. M., & Samulski, R. J. (2016). 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. *Molecular Therapy*, 24(2), 287-297. <https://doi.org/10.1038/mt.2015.187>
- Grimm, D., Kern, A., Pawlita, M., Ferrari, F. K., Samulski, R. J., & Kleinschmidt, J. A. (1999). Titration of AAV-2 particles via a novel capsid ELISA: packaging of genomes can limit production of recombinant AAV-2. *Gene Therapy*, 6(7), 1322-1330. <https://doi.org/10.1038/sj.gt.3300946>
- Han, J., Tam, K., Tam, C., Hollis, R. P., & Kohn, D. B. (2021). Improved lentiviral vector titers from a multi-gene knockout packaging line. *Molecular Therapy - Oncolytics*, 23, 582-592. <https://doi.org/10.1016/j.omto.2021.11.012>

- He, X., Urip, B. A., Zhang, Z., Ngan, C. C., & Feng, B. (2021). Evolving AAV-delivered therapeutics towards ultimate cures. *Journal of Molecular Medicine*, 99(5), 593-617. <https://doi.org/10.1007/s00109-020-02034-2>
- Hellweg, C. E., Arenz, A., Bogner, S., Schmitz, C., & Baumstark-Khan, C. (2006). Activation of Nuclear Factor κ B by Different Agents. *Annals of the New York Academy of Sciences*, 1091(1), 191-204. <https://doi.org/10.1196/annals.1378.066>
- Herzog, R. W., VandenDriessche, T., & Ozelo, M. C. (2023). First hemophilia B gene therapy approved: More than two decades in the making. *Molecular Therapy*, 31(1), 1-2. <https://doi.org/10.1016/j.ymthe.2022.12.001>
- Hidalgo, P., Anzures, L., Hernández-Mendoza, A., Guerrero, A., Wood Christopher, D., Valdés, M., Dobner, T., Gonzalez Ramón, A., & Sandri-Goldin, R. M. (2016). Morphological, Biochemical, and Functional Study of Viral Replication Compartments Isolated from Adenovirus-Infected Cells. *Journal of virology*, 90(7), 3411-3427. <https://doi.org/10.1128/JVI.00033-16>
- Hölscher, C., Hörer, M., Kleinschmidt, J. A., Zentgraf, H., Bürkle, A., & Heilbronn, R. (1994). Cell lines inducibly expressing the adeno-associated virus (AAV) rep gene: requirements for productive replication of rep-negative AAV mutants. *Journal of virology*, 68(11), 7169-7177. <https://doi.org/10.1128/JVI.68.11.7169-7177.1994>
- Hölscher, C., Kleinschmidt, J. A., & Bürkle, A. (1995). High-level expression of adeno-associated virus (AAV) Rep78 or Rep68 protein is sufficient for infectious-particle formation by a rep-negative AAV mutant. *Journal of virology*, 69(11), 6880-6885. <https://doi.org/10.1128/jvi.69.11.6880-6885.1995>
- Hösel, M., Broxtermann, M., Janicki, H., Esser, K., Arzberger, S., Hartmann, P., Gillen, S., Kleeff, J., Stabenow, D., Odenthal, M., Knolle, P., Hallek, M., Protzer, U., & Büning, H. (2012). Toll-like receptor 2-mediated innate immune response in human nonparenchymal liver cells toward adeno-associated viral vectors. *Hepatology*, 55(1), 287-297. <https://doi.org/https://doi.org/10.1002/hep.24625>
- Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44-57. <https://doi.org/10.1038/nprot.2008.211>
- Huang, M. M., & Hearing, P. (1989). The adenovirus early region 4 open reading frame 6/7 protein regulates the DNA binding activity of the cellular transcription factor, E2F, through a direct complex. *Genes & Development*, 3(11), 1699-1710. <https://doi.org/10.1101/gad.3.11.1699>
- Im, D.-S., & Muzyczka, N. (1990). The AAV origin binding protein Rep68 is an ATP-

- dependent site-specific endonuclease with DNA helicase activity. *Cell*, 61(3), 447-457. [https://doi.org/10.1016/0092-8674\(90\)90526-k](https://doi.org/10.1016/0092-8674(90)90526-k)
- Inoue, N., & Russell, D. W. (1998). Packaging Cells Based on Inducible Gene Amplification for the Production of Adeno-Associated Virus Vectors. *Journal of virology*, 72(9), 7024-7031. <https://doi.org/10.1128/jvi.72.9.7024-7031.1998>
- Jennings, K., Miyamae, T., Traister, R., Marinov, A., Katakura, S., Sowders, D., Trapnell, B., Wilson, J. M., Gao, G., & Hirsch, R. (2005). Proteasome Inhibition Enhances AAV-Mediated Transgene Expression in Human Synoviocytes in Vitro and in Vivo. *Molecular Therapy*, 11(4), 600-607. <https://doi.org/10.1016/j.ymthe.2004.10.020>
- Joiner, J., Huang, Z., McHugh, K., Stebbins, M., Aron, K., Borys, M., & Khetan, A. (2022). Process modeling of recombinant adeno-associated virus production in HEK293 cells. *Current Opinion in Chemical Engineering*, 36, 100823. <https://doi.org/10.1016/j.coche.2022.100823>
- Joshi, P. R. H., Cervera, L., Ahmed, I., Kondratov, O., Zolotukhin, S., Schrag, J., Chahal, P. S., & Kamen, A. A. (2019). Achieving High-Yield Production of Functional AAV5 Gene Delivery Vectors via Fedbatch in an Insect Cell-One Baculovirus System. *Molecular Therapy - Methods & Clinical Development*, 13, 279-289. <https://doi.org/10.1016/j.omtm.2019.02.003>
- Joshi, P. R. H., Venereo-Sanchez, A., Chahal, P. S., & Kamen, A. A. (2021). Advancements in molecular design and bioprocessing of recombinant adeno-associated virus gene delivery vectors using the insect-cell baculovirus expression platform. *Biotechnology Journal*, 16(4), 2000021. <https://doi.org/10.1002/biot.202000021>
- Jurvansuu, J., Raj, K., Stasiak, A., & Beard, P. (2005). Viral Transport of DNA Damage That Mimics a Stalled Replication Fork. *Journal of virology*, 79(1), 569-580. <https://doi.org/10.1128/JVI.79.1.569-580.2005>
- Kallunki, T., Barisic, M., Jäätelä, M., & Liu, B. (2019). How to Choose the Right Inducible Gene Expression System for Mammalian Studies? *Cells*, 8(8). <https://doi.org/10.3390/cells8080796>
- Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1), 27-30. <https://doi.org/10.1093/nar/28.1.27>
- Kay, M. A., Manno, C. S., Ragni, M. V., Larson, P. J., Couto, L. B., McClelland, A., Glader, B., Chew, A. J., J Tai, S., Herzog, R. W., Arruda, V., Johnson, F., Scallan, C., Skarsgard, E., Flake, A. W., & High, K. A. (2000). Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nature Genetics*, 24(3), 257-261. <https://doi.org/10.1038/73464>

- Kim, A., Park, J.-H., Lee, M. J., & Kim, S.-M. (2023). Interferon alpha and beta receptor 1 knockout in human embryonic kidney 293 cells enhances the production efficiency of proteins or adenoviral vectors related to type I interferons [Original Research]. *Frontiers in Bioengineering and Biotechnology*, 11. <https://doi.org/10.3389/fbioe.2023.1192291>
- King, J. A., Dubielzig, R., Grimm, D., & Kleinschmidt, J. A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *The EMBO Journal*, 20(12), 3282-3291. <https://doi.org/10.1093/emboj/20.12.3282>
- Kitada, T., DiAndreth, B., Teague, B., & Weiss, R. (2018). Programming gene and engineered-cell therapies with synthetic biology. *Science*, 359(6376), eaad1067. <https://doi.org/10.1126/science.aad1067>
- Knipe, D. M., Prichard, A., Sharma, S., & Pogliano, J. (2022). Replication Compartments of Eukaryotic and Bacterial DNA Viruses: Common Themes Between Different Domains of Host Cells. *Annual Review of Virology*, 9(1), 307-327. <https://doi.org/10.1146/annurev-virology-012822-125828>
- Kohlbrenner, E., Aslanidi, G., Nash, K., Shklyaev, S., Campbell-Thompson, M., Byrne, B. J., Snyder, R. O., Muzyczka, N., Warrington, K. H., & Zolotukhin, S. (2005). Successful Production of Pseudotyped rAAV Vectors Using a Modified Baculovirus Expression System. *Molecular Therapy*, 12(6), 1217-1225. <https://doi.org/10.1016/j.ymthe.2005.08.018>
- Kovesdi, I., & Hedley, S. J. (2010). Adenoviral Producer Cells. *Viruses*, 2(8). <https://doi.org/10.3390/v2081681>
- Kumar, R., Khandelwal, N., Thachamvally, R., Tripathi, B. N., Barua, S., Kashyap, S. K., Maherchandani, S., & Kumar, N. (2018). Role of MAPK/MNK1 signaling in virus replication. *Virus Research*, 253, 48-61. <https://doi.org/10.1016/j.virusres.2018.05.028>
- Kyung, Y.-S., Peshwa, M. V., Gryte, D. M., & Hu, W.-S. (1994). High density culture of mammalian cells with dynamic perfusion based on on-line oxygen uptake rate measurements. *Cytotechnology*, 14(3), 183-190. <https://doi.org/10.1007/BF00749615>
- Lee, Z., Lu, M., Irfanullah, E., Soukup, M., & Hu, W. S. (2022). Construction of an rAAV Producer Cell Line through Synthetic Biology. *ACS Synthetic Biology*, 11(10), 3285-3295. <https://doi.org/10.1021/acssynbio.2c00207>
- Lee, Z., Lu, M., Irfanullah, E., Soukup, M., Schmidt, D., & Hu, W. S. (2023). Development

- of an Inducible, Replication-Competent Assay Cell Line for Titration of Infectious Recombinant Adeno-Associated Virus Vectors. *Human Gene Therapy*, 34(3-4), 162-170. <https://doi.org/10.1089/hum.2022.001>
- Leppard, K. N. (1997). E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. *Journal of General Virology*, 78(9), 2131-2138. <https://doi.org/10.1099/0022-1317-78-9-2131>
- Lewis, A. M., Abu-Absi, N. R., Borys, M. C., & Li, Z. J. (2016). The use of ‘Omics technology to rationally improve industrial mammalian cell line performance. *Biotechnology and Bioengineering*, 113(1), 26-38. <https://doi.org/10.1002/bit.25673>
- Li, C., & Samulski, R. J. (2020). Engineering adeno-associated virus vectors for gene therapy. *Nature Reviews Genetics*, 21(4), 255-272. <https://doi.org/10.1038/s41576-019-0205-4>
- Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van de Peer, Y., Tavernier, J., & Callewaert, N. (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications*, 5(1), 4767. <https://doi.org/10.1038/ncomms5767>
- Lu, M., Lee, Z., Lin, Y. C., Irfanullah, I., Cai, W., & Hu, W. S. (2023). Enhancing the production of recombinant adeno-associated virus in synthetic cell lines through systematic characterization. *Biotechnology and Bioengineering*, 1-14. <https://doi.org/10.1002/bit.28562>
- Ma, Z., Ni, G., & Damania, B. (2018). Innate Sensing of DNA Virus Genomes. *Annual Review of Virology*, 5(1), 341-362. <https://doi.org/10.1146/annurev-virology-092917-043244>
- MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., & MacCoss, M. J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26(7), 966-968. <https://doi.org/10.1093/bioinformatics/btq054>
- Mainwaring, D., Joshi, A., Coronel, J., & Niehus, C. (2021). Transfer and scale-up from 10 L BioBLU® to Allegro™ STR 50 and STR 200 Bioreactors. *Cell & Gene Therapy Insights*, 7(4), 1347–1362. <https://doi.org/10.18609/cgti.2021.178>
- Manno, C. S., Pierce, G. F., Arruda, V. R., Glader, B., Ragni, M., Rasko, J. J. E., Ozelo, M. C., Hoots, K., Blatt, P., Konkle, B., Dake, M., Kaye, R., Razavi, M., Zajko, A., Zehnder, J., Rustagi, P., Nakai, H., Chew, A., Leonard, D., Wright, J. F., Lessard, R. R., Sommer, J. M., Tigges, M., Sabatino, D., Luk, A., Jiang, H., Mingozzi, F.,

- Couto, L., Ertl, H. C., High, K. A., & Kay, M. A. (2006). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nature Medicine*, 12(3), 342-347. <https://doi.org/10.1038/nm1358>
- Matsushita, T., Elliger, S., Elliger, C., Podsakoff, G., Villarreal, L., Kurtzman, G. J., Iwaki, Y., & Colosi, P. (1998). Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Therapy*, 5(7), 938-945. <https://doi.org/10.1038/sj.gt.3300680>
- Maurer, A. C., Pacouret, S., Cepeda Diaz, A. K., Blake, J., Andres-Mateos, E., & Vandenberghe, L. H. (2018). The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Nucleate Capsid Assembly. *Cell Reports*, 23(6), 1817-1830. <https://doi.org/10.1016/j.celrep.2018.04.026>
- Maurer, A. C., & Weitzman, M. D. (2020). Adeno-Associated Virus Genome Interactions Important for Vector Production and Transduction. *Human Gene Therapy*, 31(9-10), 499-511. <https://doi.org/10.1089/hum.2020.069>
- Mendell, J. R., Al-Zaidy, S., Shell, R., Arnold, W. D., Rodino-Klapac, L. R., Prior, T. W., Lowes, L., Alfano, L., Berry, K., Church, K., Kissel, J. T., Nagendran, S., L'Italien, J., Sproule, D. M., Wells, C., Cardenas, J. A., Heitzer, M. D., Kaspar, A., Corcoran, S., Braun, L., Likhite, S., Miranda, C., Meyer, K., Foust, K. D., Burghes, A. H. M., & Kaspar, B. K. (2017). Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *New England Journal of Medicine*, 377(18), 1713-1722. <https://doi.org/10.1056/NEJMoa1706198>
- Mendell, J. R., Al-Zaidy, S. A., Rodino-Klapac, L. R., Goodspeed, K., Gray, S. J., Kay, C. N., Boye, S. L., Boye, S. E., George, L. A., Salabarria, S., Corti, M., Byrne, B. J., & Tremblay, J. P. (2021). Current Clinical Applications of In Vivo Gene Therapy with AAVs. *Molecular Therapy*, 29(2), 464-488. <https://doi.org/10.1016/j.ymthe.2020.12.007>
- Merten, O.-W. (2016). AAV vector production: state of the art developments and remaining challenges. *Cell and Gene Therapy Insights*, 2(5), 521-551. <https://doi.org/10.18609/cgti.2016.067>
- Merten, O. W., Gény-Fiamma, C., & Douar, A. M. (2005). Current issues in adeno-associated viral vector production. *Gene Therapy*, 12(1), S51-S61. <https://doi.org/10.1038/sj.gt.3302615>
- Mietzsch, M., Grasse, S., Zurawski, C., Weger, S., Bennett, A., Agbandje-McKenna, M., Muzyczka, N., Zolotukhin, S., & Heilbronn, R. (2013). OneBac: Platform for Scalable and High-Titer Production of Adeno-Associated Virus Serotype 1–12 Vectors for Gene Therapy. *Human Gene Therapy*, 25(3), 212-222.

<https://doi.org/10.1089/hum.2013.184>

- Mietzsch, M., Hering, H., Hammer, E.-M., Agbandje-McKenna, M., Zolotukhin, S., & Heilbronn, R. (2017). OneBac 2.0: Sf9 Cell Lines for Production of AAV1, AAV2, and AAV8 Vectors with Minimal Encapsulation of Foreign DNA. *Human Gene Therapy Methods*, 28(1), 15-22. <https://doi.org/10.1089/hgtb.2016.164>
- Mietzsch, M., Smith, J. K., Yu, J. C., Banala, V., Emmanuel, S. N., Jose, A., Chipman, P., Bhattacharya, N., McKenna, R., & Agbandje-McKenna, M. (2020). Characterization of AAV-Specific Affinity Ligands: Consequences for Vector Purification and Development Strategies. *Molecular Therapy - Methods & Clinical Development*, 19, 362-373. <https://doi.org/10.1016/j.omtm.2020.10.001>
- Minamitani, T., Iwakiri, D., & Takada, K. (2011). Adenovirus Virus-Associated RNAs Induce Type I Interferon Expression through a RIG-I-Mediated Pathway. *Journal of virology*, 85(8), 4035-4040. <https://doi.org/doi:10.1128/JVI.02160-10>
- Mingozzi, F., & High, K. A. (2013). Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood*, 122(1), 23-36. <https://doi.org/10.1182/blood-2013-01-306647>
- Mitchell, A. M., & Samulski, R. J. (2013). Mechanistic Insights into the Enhancement of Adeno-Associated Virus Transduction by Proteasome Inhibitors. *Journal of virology*, 87(23), 13035-13041. <https://doi.org/10.1128/JVI.01826-13>
- Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E., & Hayakawa, T. (2000). IRES-Dependent Second Gene Expression Is Significantly Lower Than Cap-Dependent First Gene Expression in a Bicistronic Vector. *Molecular Therapy*, 1(4), 376-382. <https://doi.org/10.1006/mthe.2000.0050>
- Mogensen, T., H., & Paludan, S., R. (2001). Molecular Pathways in Virus-Induced Cytokine Production. *Microbiology and Molecular Biology Reviews*, 65(1), 131-150. <https://doi.org/10.1128/mnbr.65.1.131-150.2001>
- Mohiuddin, I., Loiler, S., Zolotukhin, I., Byrne, B. J., Flotte, T. R., & Snyder, R. O. (2005). Herpesvirus-based infectious titrating of recombinant adeno-associated viral vectors. *Molecular Therapy*, 11(2), 320-326. <https://doi.org/10.1016/j.ymthe.2004.08.030>
- Moreno, F., Lip, F., Rojas, H., & Anggakusuma. (2022). Development of an insect cell-based adeno-associated virus packaging cell line employing advanced Rep gene expression control system. *Molecular Therapy - Methods & Clinical Development*, 27, 391-403. <https://doi.org/10.1016/j.omtm.2022.10.015>
- Mouw, M. B., & Pintel, D. J. (2000). Adeno-associated virus RNAs appear in a temporal order and their splicing is stimulated during coinfection with adenovirus. *Journal*

- of virology*, 74(21), 9878-9888. <https://doi.org/10.1128/jvi.74.21.9878-9888.2000>
- Muhuri, M., Maeda, Y., Ma, H., Ram, S., Fitzgerald, K. A., Tai, P. W. L., & Gao, G. (2021). Overcoming innate immune barriers that impede AAV gene therapy vectors. *The Journal of clinical investigation*, 131(1). <https://doi.org/10.1172/JCI143780>
- Narni-Mancinelli, E., Gauthier, L., Baratin, M., Guia, S., Fenis, A., Deghmane, A.-E., Rossi, B., Fourquet, P., Escalière, B., Kerdiles, Y. M., Ugolini, S., Taha, M.-K., & Vivier, E. (2017). Complement factor P is a ligand for the natural killer cell-activating receptor NKp46. *Science Immunology*, 2(10), eaam9628. <https://doi.org/10.1126/sciimmunol.aam9628>
- Naso, M. F., Tomkowicz, B., Perry, W. L., & Strohl, W. R. (2017). Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs*, 31(4), 317-334. <https://doi.org/10.1007/s40259-017-0234-5>
- Oberbek, A., Matasci, M., Hacker, D. L., & Wurm, F. M. (2011). Generation of stable, high-producing cho cell lines by lentiviral vector-mediated gene transfer in serum-free suspension culture. *Biotechnology and Bioengineering*, 108(3), 600-610. <https://doi.org/https://doi.org/10.1002/bit.22968>
- Oda, Y., Okada, T., Yoshida, H., Kaufman, R. J., Nagata, K., & Mori, K. (2006). Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation. *Journal of Cell Biology*, 172(3), 383-393. <https://doi.org/10.1083/jcb.200507057>
- Ogden, P. J., Kelsic, E. D., Sinai, S., & Church, G. M. (2019). Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science*, 366(6469), 1139. <https://doi.org/10.1126/science.aaw2900>
- Ou, J., Tang, Y., Xu, J., Tucci, J., Borys, M. C., & Khetan, A. (2023). Recent advances in upstream process development for production of recombinant adeno-associated virus. *Biotechnology and Bioengineering*, n/a(n/a). <https://doi.org/10.1002/bit.28545>
- Paulk, N. (2020). Gene Therapy: It Is Time to Talk about High-Dose AAV. *Genetic Engineering & Biotechnology News*, 40(9), 14-16. <https://doi.org/10.1089/gen.40.09.04>
- Penaud-Budloo, M., François, A., Clément, N., & Ayuso, E. (2018). Pharmacology of Recombinant Adeno-associated Virus Production. *Molecular Therapy - Methods & Clinical Development*, 8, 166-180. <https://doi.org/10.1016/j.omtm.2018.01.002>
- Pereira, D. J., McCarty, D. M., & Muzyczka, N. (1997). The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription

- during a productive infection. *Journal of virology*, 71(2), 1079-1088. <https://doi.org/10.1128/JVI.71.2.1079-1088.1997>
- Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., & Coon, J. J. (2012). Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. *Molecular & Cellular Proteomics*, 11(11), 1475-1488. <https://doi.org/10.1074/mcp.O112.020131>
- Picotti, P., & Aebersold, R. (2012). Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. *Nature Methods*, 9(6), 555-566. <https://doi.org/10.1038/nmeth.2015>
- Polager, S., Kalma, Y., Berkovich, E., & Ginsberg, D. (2002). E2Fs up-regulate expression of genes involved in DNA replication, DNA repair and mitosis. *Oncogene*, 21(3), 437-446. <https://doi.org/10.1038/sj.onc.1205102>
- Prasad, V., & Greber, U. F. (2021). The endoplasmic reticulum unfolded protein response – homeostasis, cell death and evolution in virus infections. *FEMS Microbiology Reviews*, 45(5), fuab016. <https://doi.org/10.1093/femsre/fuab016>
- Qiao, C., Li, J., Skold, A., Zhang, X., & Xiao, X. (2002). Feasibility of generating adeno-associated virus packaging cell lines containing inducible adenovirus helper genes. *Journal of virology*, 76(4), 1904-1913. <https://doi.org/10.1128/jvi.76.4.1904-1913.2002>
- Qiao, C., Wang, B., Zhu, X., Li, J., & Xiao, X. (2002). A novel gene expression control system and its use in stable, high-titer 293 cell-based adeno-associated virus packaging cell lines. *Journal of virology*, 76(24), 13015-13027. <https://doi.org/10.1128/jvi.76.24.13015-13027.2002>
- Qu, W., Wang, M., Wu, Y., & Xu, R. (2015). Scalable Downstream Strategies for Purification of Recombinant Adeno- Associated Virus Vectors in Light of the Properties. *Current Pharmaceutical Biotechnology*, 16(8), 684-695. <https://doi.org/10.2174/1389201016666150505122228>
- Querido, E., Blanchette, P., Yan, Q., Kamura, T., Morrison, M., Boivin, D., Kaelin, W. G., Conaway, R. C., Conaway, J. W., & Branton, P. E. (2001). Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. *Genes & Development*, 15(23), 3104-3117. <https://doi.org/10.1101/gad.926401>
- Raj, K., Ogston, P., & Beard, P. (2001). Virus-mediated killing of cells that lack p53 activity. *Nature*, 412(6850), 914-917. <https://doi.org/10.1038/35091082>
- Rauniyar, N. (2015). Parallel Reaction Monitoring: A Targeted Experiment Performed

- Using High Resolution and High Mass Accuracy Mass Spectrometry. *International Journal of Molecular Sciences*, 16(12), 28566-28581. <https://doi.org/10.3390/ijms161226120>
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139-140. <https://doi.org/10.1093/bioinformatics/btp616>
- Rumachik, N. G., Malaker, S. A., Poweleit, N., Maynard, L. H., Adams, C. M., Leib, R. D., Cirolia, G., Thomas, D., Stamnes, S., Holt, K., Sinn, P., May, A. P., & Paulk, N. K. (2020). Methods Matter: Standard Production Platforms for Recombinant AAV Produce Chemically and Functionally Distinct Vectors. *Molecular Therapy - Methods & Clinical Development*, 18, 98-118. <https://doi.org/10.1016/j.omtm.2020.05.018>
- Russell, S., Bennett, J., Wellman, J. A., Chung, D. C., Yu, Z.-F., Tillman, A., Wittes, J., Pappas, J., Elci, O., McCague, S., Cross, D., Marshall, K. A., Walshire, J., Kehoe, T. L., Reichert, H., Davis, M., Raffini, L., George, L. A., Hudson, F. P., Dingfield, L., Zhu, X., Haller, J. A., Sohn, E. H., Mahajan, V. B., Pfeifer, W., Weckmann, M., Johnson, C., Gewaily, D., Drack, A., Stone, E., Wachtel, K., Simonelli, F., Leroy, B. P., Wright, J. F., High, K. A., & Maguire, A. M. (2017). Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *The Lancet*, 390(10097), 849-860. [https://doi.org/10.1016/S0140-6736\(17\)31868-8](https://doi.org/10.1016/S0140-6736(17)31868-8)
- Samuel, C. E. (2001). Antiviral Actions of Interferons. *Clinical Microbiology Reviews*, 14(4), 778-809. <https://doi.org/10.1128/CMR.14.4.778-809.2001>
- Samulski, R. J., & Muzyczka, N. (2014). AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. *Annual Review of Virology*, 1(1), 427-451. <https://doi.org/10.1146/annurev-virology-031413-085355>
- Saudan, P., Vlach, J., & Beard, P. (2000). Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb. *The EMBO Journal*, 19(16), 4351-4361. <https://doi.org/10.1093/emboj/19.16.4351>
- Scarrott, J. M., Johari, Y. B., Pohle, T. H., Liu, P., Mayer, A., & James, D. C. (2023). Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives. *Biotechnology Journal*, 18(3), 2200450. <https://doi.org/10.1002/biot.202200450>
- Schmidt, M., Afione, S., & Kotin Robert, M. (2000). Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. *Journal of virology*, 74(20), 9441-9450. <https://doi.org/10.1128/JVI.74.20.9441-9450.2000>

- Schwartz, R. A., Palacios, J. A., Cassell, G. D., Adam, S., Giacca, M., & Weitzman, M. D. (2007). The Mre11/Rad50/Nbs1 complex limits adeno-associated virus transduction and replication. *Journal of virology*, 81(23), 12936-12945. <https://doi.org/10.1128/JVI.01523-07>
- Shao, W., Earley, L. F., Chai, Z., Chen, X., Sun, J., He, T., Deng, M., Hirsch, M. L., Ting, J., Samulski, R. J., & Li, C. (2018). Double-stranded RNA innate immune response activation from long-term adeno-associated virus vector transduction. *JCI Insight*, 3(12). <https://doi.org/10.1172/jci.insight.120474>
- Shi, T., Song, E., Nie, S., Rodland, K. D., Liu, T., Qian, W.-J., & Smith, R. D. (2016). Advances in targeted proteomics and applications to biomedical research. *Proteomics*, 16(15-16), 2160-2182. <https://doi.org/10.1002/pmic.201500449>
- Shupe, J., Zhang, A., Odenwelder, D. C., & Dobrowsky, T. (2022). Gene therapy: challenges in cell culture scale-up. *Current Opinion in Biotechnology*, 75, 102721. <https://doi.org/10.1016/j.copbio.2022.102721>
- Smith, R. H., Levy, J. R., & Kotin, R. M. (2009). A Simplified Baculovirus-AAV Expression Vector System Coupled With One-step Affinity Purification Yields High-titer rAAV Stocks From Insect Cells. *Molecular Therapy*, 17(11), 1888-1896. <https://doi.org/10.1038/mt.2009.128>
- Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*, 136(4), 731-745. <https://doi.org/10.1016/j.cell.2009.01.042>
- Sonntag, F., Schmidt, K., & Kleinschmidt, J. A. (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(22), 10220-10225. <https://doi.org/10.1073/pnas.1001673107>
- Soriano, A. M., Crisostomo, L., Mendez, M., Graves, D., Frost, J. R., Olanubi, O., Whyte, P. F., Haring, P., & Pelka, P. (2019). Adenovirus 5 E1A Interacts with E4orf3 To Regulate Viral Chromatin Organization. *Journal of virology*, 93(10), e00157-00119. <https://doi.org/10.1128/JVI.00157-19>
- Srivastava, A. (2016). In vivo tissue-tropism of adeno-associated viral vectors. *Current opinion in virology*, 21, 75-80. <https://doi.org/10.1016/j.coviro.2016.08.003>
- Srivastava, A., Lusby, E. W., & Berns, K. I. (1983). Nucleotide sequence and organization of the adeno-associated virus 2 genome. *Journal of virology*, 45(2), 555-564. <https://doi.org/10.1128/JVI.45.2.555-564.1983>

- Srivastava, A., Mallela, K. M. G., Deorkar, N., & Brophy, G. (2021). Manufacturing Challenges and Rational Formulation Development for AAV Viral Vectors. *Journal of Pharmaceutical Sciences*, 110(7), 2609-2624. <https://doi.org/10.1016/j.xphs.2021.03.024>
- Strasser, L., Boi, S., Guapo, F., Donohue, N., Barron, N., Rainbow-Fletcher, A., & Bones, J. (2021). Proteomic Landscape of Adeno-Associated Virus (AAV)-Producing HEK293 Cells. *International Journal of Molecular Sciences*, 22(21), 11499. <https://doi.org/10.3390/ijms222111499>
- Strobel, B., Klauser, B., Hartig, J. S., Lamla, T., Gantner, F., & Kreuz, S. (2015). Riboswitch-mediated Attenuation of Transgene Cytotoxicity Increases Adeno-associated Virus Vector Yields in HEK-293 Cells. *Molecular Therapy*, 23(10), 1582-1591. <https://doi.org/10.1038/mt.2015.123>
- Stutika, C., Gogol-Döring, A., Botschen, L., Mietzsch, M., Weger, S., Feldkamp, M., Chen, W., Heilbronn, R., & Banks, L. (2016). A Comprehensive RNA Sequencing Analysis of the Adeno-Associated Virus (AAV) Type 2 Transcriptome Reveals Novel AAV Transcripts, Splice Variants, and Derived Proteins. *Journal of virology*, 90(3), 1278-1289. <https://doi.org/10.1128/JVI.02750-15>
- Su, W., Patrício, M. I., Duffy, M. R., Krakowiak, J. M., Seymour, L. W., & Cawood, R. (2022). Self-attenuating adenovirus enables production of recombinant adeno-associated virus for high manufacturing yield without contamination. *Nature Communications*, 13(1), 1182. <https://doi.org/10.1038/s41467-022-28738-2>
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43), 15545-15550. <https://doi.org/10.1073/pnas.0506580102>
- Subramanian, S., Maurer, A. C., Bator, C. M., Makhov, A. M., Conway, J. F., Turner, K. B., Marden, J. H., Vandenberghe, L. H., & Hafenstein, S. L. (2019). Filling Adeno-Associated Virus Capsids: Estimating Success by Cryo-Electron Microscopy. *Human Gene Therapy*, 30(12), 1449-1460. <https://doi.org/10.1089/hum.2019.041>
- Sun, J. Y., Anand-Jawa, V., Chatterjee, S., & Wong, K. K. (2003). Immune responses to adeno-associated virus and its recombinant vectors. *Gene Therapy*, 10(11), 964-976. <https://doi.org/10.1038/sj.gt.3302039>
- Tang, Q., Keeler, A. M., Zhang, S., Su, Q., Lyu, Z., Cheng, Y., Gao, G., & Flotte, T. R. (2020). Two-Plasmid Packaging System for Recombinant Adeno-Associated Virus. *BioResearch Open Access*, 9(1), 219-228. <https://doi.org/10.1089/biores.2020.0031>

- Tavassoly, I., Goldfarb, J., & Iyengar, R. (2018). Systems biology primer: the basic methods and approaches. *Essays in Biochemistry*, 62(4), 487-500. <https://doi.org/10.1042/EBC20180003>
- Tenenbaum, L., Chtarto, A., Lehtonen, E., Velu, T., Brotchi, J., & Levivier, M. (2004). Recombinant AAV-mediated gene delivery to the central nervous system. *The Journal of Gene Medicine*, 6(S1), S212-S222. <https://doi.org/10.1002/jgm.506>
- Thomas, D. L., Wang, L., Niamke, J., Liu, J., Kang, W., Scotti, M. M., Ye, G.-J., Veres, G., & Knop, D. R. (2009). Scalable Recombinant Adeno-Associated Virus Production Using Recombinant Herpes Simplex Virus Type 1 Coinfection of Suspension-Adapted Mammalian Cells. *Human Gene Therapy*, 20(8), 861-870. <https://doi.org/10.1089/hum.2009.004>
- Tipanee, J., VandenDriessche, T., & Chuah, M. K. (2017). Transposons: Moving Forward from Preclinical Studies to Clinical Trials. *Human Gene Therapy*, 28(11), 1087-1104. <https://doi.org/10.1089/hum.2017.128>
- Tomono, T., Hirai, Y., Okada, H., Miyagawa, Y., Adachi, K., Sakamoto, S., Kawano, Y., Chono, H., Mineno, J., Ishii, A., Shimada, T., Onodera, M., Tamaoka, A., & Okada, T. (2018). Highly Efficient Ultracentrifugation-free Chromatographic Purification of Recombinant AAV Serotype 9. *Molecular Therapy - Methods & Clinical Development*, 11, 180-190. <https://doi.org/10.1016/j.omtm.2018.10.015>
- Trempe, J. P., & Carter, B. J. (1988). Regulation of adeno-associated virus gene expression in 293 cells: control of mRNA abundance and translation. *Journal of virology*, 62(1), 68-74. <https://doi.org/10.1128/jvi.62.1.68-74.1988>
- Tullis Gregory, E., & Shenk, T. (2000). Efficient Replication of Adeno-Associated Virus Type 2 Vectors: a cis-Acting Element outside of the Terminal Repeats and a Minimal Size. *Journal of virology*, 74(24), 11511-11521. <https://doi.org/10.1128/JVI.74.24.11511-11521.2000>
- Urabe, M., Ding, C., & Kotin, R. M. (2002). Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. *Human Gene Therapy*, 13(16), 1935-1943. <https://doi.org/10.1089/10430340260355347>
- Vachon, V. K., & Conn, G. L. (2016). Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Research*, 212, 39-52. <https://doi.org/10.1016/j.virusres.2015.06.018>
- van Lieshout, L. P., Rubin, M., Costa-Grant, K., Ota, S., Golebiowski, D., Panico, T., Wiberg, E., Szymczak, K., Gilmore, R., Stanvick, M., Burnham, B., Gagnon, J., Iwuchukwu, I., Yang, G., Ghazi, I., Meola, A., Dickerson, R., Thiers, T., Mustich,

- L., Hayes, A., Rivas, I., Lotterhand, J., Avila, N., McGivney, J., Yin, J., & Kelly, T. (2023). A novel dual-plasmid platform provides scalable transfection yielding improved productivity and packaging across multiple AAV serotypes and genomes. *Molecular Therapy - Methods & Clinical Development*, 29, 426-436. <https://doi.org/10.1016/j.omtm.2023.05.004>
- Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, 13(4), 227-232. <https://doi.org/10.1038/nrg3185>
- Wang, B., Li, J., Fu, F. H., Chen, C., Zhu, X., Zhou, L., Jiang, X., & Xiao, X. (2008). Construction and analysis of compact muscle-specific promoters for AAV vectors. *Gene Therapy*, 15(22), 1489-1499. <https://doi.org/10.1038/gt.2008.104>
- Wang, Y., Fu, Q., Lee, Y. S., Sha, S., & Yoon, S. (2023). Transcriptomic features reveal molecular signatures associated with recombinant adeno-associated virus production in HEK293 cells. *Biotechnology Progress*, 39(4), e3346. <https://doi.org/10.1002/btpr.3346>
- Ward, P., Dean, F. B., O'Donnell, M. E., & Berns, K. I. (1998). Role of the Adenovirus DNA-Binding Protein in In Vitro Adeno-Associated Virus DNA Replication. *Journal of virology*, 72(1), 420-427. <https://doi.org/10.1128/JVI.72.1.420-427.1998>
- Weitzman, M. D., Fisher, K. J., & Wilson, J. M. (1996). Recruitment of wild-type and recombinant adeno-associated virus into adenovirus replication centers. *Journal of virology*, 70(3), 1845-1854. <https://doi.org/10.1128/jvi.70.3.1845-1854.1996>
- Weitzman, M. D., & Fradet-Turcotte, A. (2018). Virus DNA Replication and the Host DNA Damage Response. *Annual Review of Virology*, 5(1), 141-164. <https://doi.org/10.1146/annurev-virology-092917-043534>
- Wright, J. F. (2008). Manufacturing and characterizing AAV-based vectors for use in clinical studies. *Gene Therapy*, 15(11), 840-848. <https://doi.org/10.1038/gt.2008.65>
- Wright, J. F. (2014). Product-Related Impurities in Clinical-Grade Recombinant AAV Vectors: Characterization and Risk Assessment. *Biomedicines*, 2(1), 80-97. <https://doi.org/10.3390/biomedicines2010080>
- Wu, Z., Yang, H., & Colosi, P. (2010). Effect of Genome Size on AAV Vector Packaging. *Molecular Therapy*, 18(1), 80-86. <https://doi.org/10.1038/mt.2009.255>
- Wyborski, D. L., & Short, J. M. (1991). Analysis of inducers of the E.coli lac repressor system mammalian cells and whole animals. *Nucleic Acids Research*, 19(17), 4647-

4653. <https://doi.org/10.1093/nar/19.17.4647>

- Xiao, X., Li, J., & Samulski, R. J. (1996). Efficient long-term gene transfer into muscle tissue of immunocompetent mice by adeno-associated virus vector. *Journal of virology*, 70(11), 8098-8108. <https://doi.org/10.1128/Jvi.70.11.8098-8108.1996>
- Xiao, X., Li, J., & Samulski, R. J. (1998). Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. *Journal of virology*, 72(3), 2224. <https://doi.org/10.1128/JVI.72.3.2224-2232.1998>
- Yang, L., Gilbertsen, A., Smith, K., Xia, H., Higgins, L., Guerrero, C., & Henke, C. A. (2022). Proteomic analysis of the IPF mesenchymal progenitor cell nuclear proteome identifies abnormalities in key nodal proteins that underlie their fibrogenic phenotype. *Proteomics*, 22(13-14), 2200018. <https://doi.org/10.1002/pmic.202200018>
- Yang, Q., Chen, F., & Trempe, J. P. (1994). Characterization of cell lines that inducibly express the adeno-associated virus Rep proteins. *Journal of virology*, 68(8), 4847-4856. <https://doi.org/10.1128/jvi.68.8.4847-4856.1994>
- Ye, Q., Phan, T., Hu, W.-S., Liu, X., Fan, L., Tan, W.-S., & Zhao, L. (2021). Transcriptomic Characterization Reveals Attributes of High Influenza Virus Productivity in MDCK Cells. *Viruses*, 13(11), 2200. <https://doi.org/10.3390/v13112200>
- Zaiss, A., K., Cotter, M., J., White, L., R., Clark, S., A., Wong, N., C. W., Holers, V. Michael, Bartlett, J., S., & Muruve, D., A. (2008). Complement Is an Essential Component of the Immune Response to Adeno-Associated Virus Vectors. *Journal of virology*, 82(6), 2727-2740. <https://doi.org/10.1128/JVI.01990-07>
- Zarubin, T., & Han, J. (2005). Activation and signaling of the p38 MAP kinase pathway. *Cell Research*, 15(1), 11-18. <https://doi.org/10.1038/sj.cr.7290257>
- Zeltner, N., Kohlbrenner, E., Clement, N., Weber, T., & Linden, R. M. (2010). Near-perfect infectivity of wild-type AAV as benchmark for infectivity of recombinant AAV vectors. *Gene Therapy*, 17(7), 872-879. <https://doi.org/10.1038/gt.2010.27>
- Zhang, L., Inniss, M. C., Han, S., Moffat, M., Jones, H., Zhang, B., Cox, W. L., Rance, J. R., & Young, R. J. (2015). Recombinase-mediated cassette exchange (RMCE) for monoclonal antibody expression in the commercially relevant CHOK1SV cell line. *Biotechnology Progress*, 31(6), 1645-1656. <https://doi.org/10.1002/btpr.2175>
- Zhang, W., Xiao, D., Shan, L., Zhao, J., Mao, Q., & Xia, H. (2017). Generation of apoptosis-resistant HEK293 cells with CRISPR/Cas mediated quadruple gene knockout for improved protein and virus production. *Biotechnology and*

Bioengineering, 114(11), 2539-2549. <https://doi.org/10.1002/bit.26382>

- Zhang, X., Alwis, M. D., Hart, S. L., Fitzke, F. W., Inglis, S. C., Bournsnel, M. E. G., Levinsky, R. J., Kinnon, C., Ali, R. R., & Thrasher, A. J. (1999). High-Titer Recombinant Adeno-Associated Virus Production from Replicating Amplicons and Herpes Vectors Deleted for Glycoprotein H. *Human Gene Therapy*, 10(15), 2527-2537. <https://doi.org/10.1089/10430349950016861>
- Zhao, H., Lee, K.-J., Daris, M., Lin, Y., Wolfe, T., Sheng, J., Plewa, C., Wang, S., & Meisen, W. H. (2020). Creation of a High-Yield AAV Vector Production Platform in Suspension Cells Using a Design-of-Experiment Approach. *Molecular Therapy - Methods & Clinical Development*, 18, 312-320. <https://doi.org/10.1016/j.omtm.2020.06.004>
- Zhou, X., Zolotukhin, I., Im, D.-S., & Muzyczka, N. (1999). Biochemical Characterization of Adeno-Associated Virus Rep68 DNA Helicase and ATPase Activities. *Journal of virology*, 73(2), 1580-1590. <https://doi.org/10.1128/jvi.73.2.1580-1590.1999>
- Zhu, J., Huang, X., & Yang, Y. (2009). The TLR9-MyD88 pathway is critical for adaptive immune responses to adeno-associated virus gene therapy vectors in mice. *The Journal of clinical investigation*, 119(8), 2388-2398. <https://doi.org/10.1172/JCI37607>