4-30-2018

Optimization of Recombinant Adeno-Associated Virus (aav) Vector Production in Saccharomyces Cerevisiae

Juan Aponte

OPTIMIZATION OF RECOMBINANT ADENO-ASSOCIATED VIRUS (AAV) VECTOR PRODUCTION IN *SACCHAROMYCES CEREVISIAE*

BY

JUAN JOSE APONTE-UBILLUS

A Dissertation submitted to the Faculty of Keck Graduate Institute of Applied Life Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Applied Life Sciences.

Claremont, California
2018

Approved by:

DocuSigned by:
Parviz Shamlou
6D5CE8DB65CE453...

Parviz Shamlou, Chair

We, the undersigned, certify that we have read this dissertation of Juan Jose Aponte-Ubillus and approve it as adequate in scope and quality for the degree of Doctor of Philosophy.

Dissertation Committee:

DocuSigned by:
Parviz Shamlou
6D5CE8DB65CE453...

Parviz Shamlou, Chair

DocuSigned by:
Joseph Peltier
3558C982F4A3421...

Joe Peltier, Member

DocuSigned by:
Daniel Barajas
73922094E03F42C...

Daniel Barajas, Member

DocuSigned by:
Cameron Bardliving
67C81C8E090E40B...

Cameron Bardliving, Member, PhD Program Director

Optimization of recombinant adeno-associated virus (rAAV) vector production in *Saccharomyces cerevisiae*

By:

Juan Aponte-Ubillus

Keck Graduate Institute, 2018

ABSTRACT

Recombinant adeno-associated viral vectors (rAAV) are emerging drugs for gene therapy applications. Their non-pathogenic status, low inflammatory potential, availability of viral serotypes with different tissue tropisms, and prospective long-lasting gene expression are important attributes that make rAAVs safe and efficient therapeutic options. One of the main limitations for bringing rAAV gene therapy to the market is the difficulty in supplying enough rAAV vector product. The high vector doses suggested by early clinical data infer the need to scale up production at high volumes in order to satisfy patient demand. Current production platforms such as HEK293 or Sf9 cells are very efficient but up to date, scalability issues limit their use to preclinical and phase I/II production campaigns. Our team recently developed a novel rAAV-producing yeast strain, which recapitulated key molecular processes for vector particle formation (Barajas et al., 2017). The use of a microbial system for vector production would represent an affordable and highly scalable platform for industrial production. Preliminary data showed low vector yields, possibly associated to very low DNA encapsidation rate.

The present thesis aims at getting clues about the molecular and bioprocessing factors that could be impacting vector yield in the novel rAAV-producing yeast system. In one approach, we performed a proteomic profiling of the yeast host response to rAAV protein expression. By using mass spectrometry and bioinformatics tools, we were able to identify trends in protein expression associated to vector formation. Gene ontology enrichment and network interaction analyses highlighted five specific cellular events: protein folding activity linked to unfolded protein response, proteasomal degradation activity, oxidation-reduction processes linked to oxidative stress, protein biosynthesis, and carbon metabolism. We speculated that some of these processes might be directly or indirectly linked to vector production constraints. A protein overexpression strategy was tested by transforming yeast with 2-micron plasmids carrying expression cassettes for 19 host cell proteins identified in the profiling. Increased vector yield was obtained in yeast strains overexpressing proteins SSA2, SSE1, SSE2, CCP1, GTT1, and GAL4.

On a second approach, we used the yeast system as a means to screen the effect of host protein expression modulation on rAAV DNA replication and vector yield, by using the yTHC library strains (R1158-derived) and a set of 2 plasmids that confer all rAAV genetic elements. More than 850 strains, each one with a single host gene under a TET-repressible promoter, were screened in duplicates. From preliminary screenings, we identified 22 gene candidates that improved rAAV DNA replication (rAAV-GFP/18s rDNA ratio) and vector yield (benzonase-resistant rAAV DNA vector genome titer) as high as 6-fold and 15-fold relative to control, respectively. The candidate proteins participate in various biological processes such as DNA replication, ribosome biogenesis, and RNA and protein processing. The top five candidates (*PRE4, HEM4, TOP2, GPN3, SDO1*) were further screened by generating overexpression mutants in another yeast strain (YPH500). Subsequent clone evaluation was performed to confirm the rAAV-promoting activity of

selected candidates under plate-based and bioreactor-controlled fermentation conditions. Our results highlighted HEM4 and TOP2 proteins as enhancers of rAAV2 vector yield in the yeast model.

A final approach was focused on bioprocessing studies intended to develop a fed-batch fermentation process for rAAV2 vector production. Preliminary characterization studies performed in shake flasks provided useful data regarding rAAV DNA replication and vector formation in yeast over time, as well as optimal pH and temperature values for fermentation. Results suggested extending the original process to four days of galactose induction, and operating values of starting pH and temperature of 4.8 and 30°C, respectively. An additional media optimization study was performed to identify critical media components for optimal vector yield. A 3-fold increase was obtained after supplementing the galactose induction media with lysine, pyridoxin, myo-inositol, ferric chloride, and cysteine. We were able to translate a shake flask-based, batch process with medium replacement to a bioreactor-controlled fed-batch process. Low and moderate cell culture strategies were performed, controlling pH, DO %, and temperature. Additional studies were done to optimize growth rate, glucose and galactose feed, and induction strategy. However, final yields at moderate cell densities were comparable to the ones obtained at low cell densities, suggesting the presence of unknown factors that might be impacting per cell productivity.

These three independent approaches provided important information regarding molecular and process strategies to optimize rAAV vector yield. Follow-up studies need to be done to consolidate yeast strain development and fermentation development efforts into a robust yeast platform potentially useful for industrial vector production.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Joe Peltier and Daniel Barajas at Biomarin Pharmaceutical. Thank you for giving me your vote of confidence to command this short and exciting investigation. There are no words to thank you for your time, patience, guidance and mentoring throughout the execution of the project, and I look forward to continuing my contribution to your teams, now as a member of the scientific staff. I would also like to thank Dan Gold for being a big supporter of technology development; and to thank all the members of the Process Sciences department at Biomarin Pharmaceutical, for sharing words of encouragement and good vibes with me over the last 3 years.

I would like to thank Parviz Shamlou and Cameron Bardliving at Keck Graduate Institute. Their advice as experienced bioprocess engineers gave this project a good balance between strong basic research and applied life sciences and engineering. Thank you for your willingness to collaborate on this offsite project and to be always ready to comment, argue or disagree with the content. Your feedback was instrumental at difficult times, and help me stay focus on the practical objectives of the project. I would like to extend this acknowledgement to KGI faculty and staff, especially to Sofia Toro, Elizabeth Wright, Illya Tolstorukov, Ian Phillips and Kirilynn Svay; who contributed in many different ways to make my stay at the school really enjoyable.

Last but not least, I thank my parents for their constant support since I started this long journey; I thank my wife Lily for her love, patience and understanding while I was doing full-time student researcher duties; and to God Almighty who is always with me and makes sure I get surrounded by good people during life.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

rAAV	recombinant adeno-associated virus
SC	synthetic complete
SCID	severe combined immunodeficiency
LPL	lipoprotein lipase
OTC	ornithine transcarbamylase
AdV	adenovirus
HSV	herpes simplex virus
BV	baculovirus
BHK	baby hamster kidney
HEK	human embryonic kidney
Sf	Spodoptera furgiperda
MOI	multiplicity of infection
GRAS	generally recongnized as safe
ORF	open reading frame
ITR	inverted terminal repeats
AAP	assembly activating protein
PLA2	phospholipase A2
NLS	nuclear localization signal
MCM	minichromosome maintenance

RFA	replication factor A
PCNA	proliferating cell nuclear antigen
WT	wild-type
Rc	replication-competent
Vg	vector genome
AcMNPV	*Autographa californica* nucleopolyhedrovirus
VLP	virus-like particles
IE1	immediate early 1 promoter
Polh	polyhedron promoter
DIP	defective interfering particles
RBE	Rep binding element
YB1	Y box element 1
HPV	human papilloma virus
TBSV	tomato bushy stunt virus
IMBYMV	Indian mung bean yellow mosaic begomovirus
GFP	green fluorescent protein
GOI	gene of interest
MMLV	Moloney murine leukemia virus
DoE	design of experiments
YPD	yeast extract-peptone-dextrose medium

yTHC	yeast TET Hughes collection
RQ	respiratory quotient
DO%	dissolved oxygen percentage
Dox	doxycycline
PEG	polyethylene glycol
LiAc	lithium acetate
MS	mass spectrometry
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride powder
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
UPLC	Ultra performance liquid chromatography
PCA	principal component analysis
ER	endoplasmic reticulum
Hsp	heat shock protein
KOH	potassium hydroxide
cfu	colony-forming unit

LIST OF FIGURES

Chapter 2

Chapter 4

Chapter 5

Chapter 6

LIST OF TABLES

Chapter 2

Chapter 5

Chapter 6

SUPPLEMENTARY INFORMATION

Chapter 4

Chapter 5

Chapter 6

CHAPTER 1. INTRODUCTION

Gene Therapy

Gene therapy comprises a group of techniques that delivers nucleic acids into patient's cells in order to correct genetic deficiencies. This is usually accomplished by the addition, correction or silencing of a target gene. Up to date, more than 2,000 gene therapy clinical trials have been carried out for a variety of medical conditions such as cancer, cardiovascular and monogenetic diseases (Gene therapy Clinical Trials Worldwide, Wiley 2017). Based on the gene delivery method, therapies are classified as non-viral or viral. The first method includes the use of liposomes, cationic complexes or sophisticated tools like a gene gun to deliver nucleic acids; whereas the latter method requires the use of viral vectors with strong transduction efficiency and tissue specificity. A different classification for gene therapy methods based on the location of the target cell during gene delivery refer to them as *ex-vivo* (target cells are harvested from the patient's body before gene therapy) or *in-vivo* (target cells get transduced at their natural anatomical location, inside the patient's body).

Surveys on gene therapy clinical trials over the last decades show a steady increase in the number of viral gene therapy clinical studies, a fact that reflects the momentum gained by this emerging group of therapies. Lentivirus, retrovirus and adenovirus remain as the most used vectors for these studies. However, the number of studies that utilizes alternative vectors (e.g, Herpesvirus, Vaccinia virus, Adeno-associated virus) has increased over the last decade (Gene therapy Clinical Trials Worldwide, Wiley 2017).

From a historical perspective, encouraging preclinical results set the path for the first gene therapy studies in humans. A lot of attention was focused on severe combined immunodeficiency (SCID), also known as the bubble boy syndrome. Many *ex-vivo* therapies were attempted, with initial studies showing inconclusive results. In 2000, the first successful SCID clinical trial was reported (Retrovirus-based therapy for SCID-X1) (Sheridan 2011). The success of this study was based on the therapeutic levels achieved during the study. There were still considerations to be accounted for to prevent inflammatory reactions and potential insertional mutagenesis.

In 1999, the first death during a gene therapy trial was reported. Jesse Gelsinger died receiving treatment for Ornithine Transcarbamylase (OTC) deficiency, in a controversial case of vector-associated toxicity (Sheridan 2011). Press releases associated to this event were spread around the international community and unfortunately it had strong repercussions in future development of trials. Nevertheless, the field was able to overcome this hurdle after several years. In 2012, a company gained market authorization for the first gene therapy product. Uniqure received EMA approval for Glybera, an AAV-based therapy for lipoprotein lipase (LPL) deficiency. In 2017, FDA granted market approval to Novartis for Kymriah, a novel CAR-T cell therapy for the treatment of leukemia. Recently, Spark therapeutics gained FDA approval for Luxturna, an AAV-based therapy intended to treat a congenital form of retinal dystrophy.

Recombinant adeno-associated virus (rAAV) as vector for in-vivo gene therapy

The use of AAV as delivery carrier for gene therapy constitutes an attractive option for the pharmaceutical industry. Recombinant AAV is clinically recognized as non-pathogenic; in addition, it has shown low inflammatory response derived from its interaction with the

human body (Kay 2011). These two characteristics put rAAV above other vectors such as herpesvirus and adenovirus, which tend to drive a strong immune response. Secondly, the very low rate of AAV vector genetic integration reduces any risk of potential insertional mutagenesis. Other advantageous biological features of rAAV include the availability of multiple serotypes for targeting a variety of organs, and the potential for long term gene expression.

rAAV vector production.- Main features and limitations

There are three main production platforms currently used for AAV production, which are based on the use of different complementation viruses. The adenovirus complementation system receives its name because of the need of infecting the host cell line with adenovirus to generate rAAV. One of the first generations of this platform comprises a double transfection process (Rep/Cap DNA and transgene DNA plasmids) and the successive addition of wild type Adenovirus. Once these three components are added to HEK293 cell line, VP proteins are produced and Rep proteins are expressed, triggering rAAV DNA replication and packaging. The identification of the minimal set of adenovirus helper genes permitted the generation of a third plasmid, replacing the use of wild type adenovirus. This triple transfection method is the most commonly used platform for the generation of pre-clinical and clinical vector material. Although this method is very efficient for rAAV production, it shows some disadvantages that limit its use for commercial production, process scalability and potential encapsidation of host cell oncogenes being the main concerns (Ayuso 2016). Many vector core facilities around the country use adherent cell lines for rAAV production. Performing scale-up based on surface area (T-flask or roller bottle cultures) becomes an extremely laborious, expensive and time-consuming task. Increasing dosing requirements for gene therapy could require

bioproduction in very large tanks (multiple thousands of liters), thus limiting the applicability of anchorage-dependent cultures to small scale production.

The herpesvirus complementation system constitutes another rAAV production platform. Although different mammalian cell lines could be used as a host for production, HSV infection seems to occur more efficiently in BHK cells. This process comprises an initial viral infection with two recombinant, replication-deficient HSV virus, one containing Rep/Cap genes and the other one containing the transgene. BHK cells adapt well to suspension culture and have very competitive per-cell productivity values, making it an attractive platform. However, the main disadvantages of this system are the genetic instability of the viral inoculum and its sensitivity to processing conditions (Aponte-Ubillus et al. 2018).

A third platform relies on the use of the insect cell / Baculovirus (BV) expression system. This vector production platform was initially described by Urabe and collaborators, as a potential solution to overcome low volumetric productivities from adherent cell-based processes. The Sf9 insect cell line is commonly used for this process, however other cell lines such as Sf21 and High Five could be utilized as well. The first-generation process required cell infection with three different recombinant Baculoviruses (Rep, Cap and Transgene) at relatively high multiplicity of infection (MOI). Further molecular biology improvements and generation of packaging and producer cell lines allowed for the reduction of required rBV strains, increasing the robustness of the process and lowering the MOI requirements. Overall, the BV system constitutes one of the most productive AAV platforms up to date.

Production of clinical-grade AAV vector has been accomplished with the use of these three AAV-producing platforms. Unfortunately, process scalability issues limit their use to preclinical and phase I/II production campaigns. Increasing vector dose requirements justifies the need for highly scalable processes that can meet potential commercial demand. Therefore, new processes will need to be developed for future phase III/commercial production of AAV vectors. Exploring microbial production of AAV vectors might constitute an attractive option.

Saccharomyces as a potential platform for rAAV production

Saccharomyces constitutes a simple eukaryotic expression system, capable of producing heterologous proteins, antigens, and virus-like particles. The use of *Saccharomyces* as a platform for recombinant production has many advantages:

1. Fully-sequenced genome
2. Simple nutritional requirements
3. Short doubling time in culture
4. Considered by FDA as Generally Recognized As Safe (GRAS)
5. Amenable to very large scale fermentation processes
6. Plethora of constitutive and inducible promoters
7. Many genomic, proteomic and bioinformatics tools available

As a production system, *Saccharomyces* has a proven record of successful vaccine production, including the manufacturing of specific viral components as monomers (primarily viral capsid antigens) and as complex layered particles (virus-like particles). Commercial products made in yeast include Euvax-B (LG Life Sciences), Engerix-B (GSK), Recombivax-HB (Merck), and Gardasil (Merck).

In addition, *Saccharomyces* represents an excellent model to study virus-host interactions. Previous efforts reported the development of yeast models that support viral DNA/RNA replication (Alves-Rodrigues et al. 2006). These models constitute helpful tools to study the biology of human and plant pathogens such as Human Papillomavirus and Tomato Bushy Stunt Virus, respectively.

Recent literature has evidenced successful replication of AAV transgene DNA and formation of infectious viral particles in yeast (Barajas et al. 2017). However, low vector yields are obtained at the end of the fermentation process. More studies are needed to get a better understanding of the main biological processes driving vector production and to identify production constraints.

Aim and scope of this research

Adeno-associated viruses (AAV) are considered excellent delivery systems for gene therapy. Within the biotech industry, animal-based AAV-producing hosts have been established. However, these platforms face technical challenges that limit their applicability for commercial vector production. From a molecular biology-based perspective, viral factors associated to vector production have been extensively studied. However, the implication of host cell factors in this phenomenon is not well understood. From an engineering-based perspective, there is a need to study and develop highly scalable and productive platforms for the biomanufacturing of gene therapy vectors that can satisfy phase III and commercial drug production requirements.

Due to its easy scalability, simple media requirements, proven safety, and process robustness, production of AAV in yeast could potentially meet large scale requirements. In particular, *Saccharomyces* genetics and metabolism are very well-understood, making

it an ideal platform model for studying the role of host-cell proteins in AAV replication and packaging. From that standpoint, we have generated a new platform for recombinant AAV production based on the baker's yeast, *Saccharomyces cerevisiae.* Preliminary results demonstrate the capacity of this model organism to produce infectious AAV particles. An in-depth study of the capabilities of this host is essential for the understanding of biological and physicochemical parameters that exert an effect on AAV replication, assembly and packaging. The scope of this research will focus on the molecular and process design of an AAV-producing yeast platform. Specifically, my first objective focuses on the identification of host cell factors that influence vector production in yeast, working towards the generation of an enhanced cell line. My second objective is to understand the influence of bioprocessing parameters on this new biological system and design a fermentation process for AAV vector production.

Significance of this research

Our *S. cerevisiae* strain represents an eukaryotic study model for rAAV production. An in-depth genetic analysis of this well characterized host will bring us one step closer to the understanding of the mechanisms underlying rAAV replication, assembly and packaging. In addition, the potential identification of enhancing host cell factors in yeast could be translated into further studies of genetic analogs in other eukaryotic cell lines like Sf9 or HEK293.

In addition, this novel yeast system could potentially become a rAAV-producing platform. Further development will be instrumental for the evaluation of its productivity and assessment of its potential implementation at industrial scale.

Overview of the study

The background chapter focuses on foundations of AAV biology and vector production. It provides a critical review of the main rAAV-producing platforms, covering recent molecular biology and engineering work, as well as limitations for large-scale manufacturing. It is in this scenario where the concept of microbial production of rAAV comes to place. This chapter includes a final section in *Saccharomyces* basic biology and a review of its utility as a model for virus DNA replication, and as a host for production of heterologous viral proteins.

In the next chapter, the research objectives for this thesis project are defined. The mechanisms underlying rAAV production in yeast are not fully understood, and that lack of knowledge plus the possibility of developing a new vector production system are driving forces for this research. My research questions are oriented towards the identification of cellular components and physicochemical factors that influence AAV production in yeast.

The experimental plan designed to answer my research questions is covered in chapters four, five and six. The first two chapters detail genomic and proteomic approaches to profile protein diversity during induction of rAAV production, and screen the effect of host cell protein overexpression on rAAV DNA replication and vector yield. The main goal of chapters four and five is to identify proteins that impact rAAV production; this information ultimately would help us establishing an enhanced cell line.

Chapter six comprises efforts to develop a small scale fed-batch fermentation process for rAAV production. This section explores the impact of physicochemical parameters on the growth and metabolic behavior of yeast.

References

1. Alves-Rodrigues I, Galão RP, Meyerhans A, Díez J (2006) Saccharomyces cerevisiae: a useful model host to study fundamental biology of viral replication. Virus Res 120:49–56 . doi: 10.1016/j.virusres.2005.11.018
2. Aponte-Ubillus JJ, Barajas D, Peltier J, Bardliving C, Shamlou P, Gold D (2017) Molecular design for recombinant adeno-associated virus (rAAV) vector production. Applied Microbiology and Biotechnology. doi: 10.1007/s00253-017-8670-1
3. Ayuso E (2016) Manufacturing of recombinant adeno-associated viral vectors: new technologies are welcome. Mol Ther Methods Clin Dev 3:15049 . doi: 10.1038/mtm.2015.49
4. Barajas D, Aponte-Ubillus JJ, Akeefe H, Cinek T, Peltier J, Gold D (2017) Generation of infectious recombinant Adeno-associated virus in Saccharomyces cerevisiae. PLoS ONE 12:e0173010 . doi: 10.1371/journal.pone.0173010
5. Kay MA (2011) State-of-the-art gene-based therapies: the road ahead. Nat Rev Genet 12:316–328 . doi: 10.1038/nrg2971
6. Sheridan C (2011) Gene therapy finds its niche. Nat Biotechnol 29:121–128 . doi: 10.1038/nbt.1769
7. Gene Therapy Clinical Trials Worldwide. (2016) http://www.wiley.com//legacy/wileychi/genmed/clinical/. Accessed 24 Nov 2016

CHAPTER 2. LITERATURE REVIEW

Adeno-associated virus biology

Physical and genetic structure

Adeno-associated virus (AAV) is a non-enveloped, single-stranded DNA virus that belongs to the family *Parvoviridae*, genus *Dependoparvovirus*. The AAV virion consists of a 25nm-diameter icosahedral capsid and a 4.7 kb linear DNA genome flanked by lateral T-shaped inverted terminal repeats (ITRs) (Samulski et al. 2014). The genome contains 3 open reading frames (orf): Rep orf encodes for four regulatory proteins that play important roles in replication and encapsidation of AAV DNA (Rep78, Rep68, Rep52, and Rep40); Cap orf encodes for three capsid proteins (VP1, VP2, VP3); and the third orf encodes for an AAV assembly activating protein (AAP) (Goncalves 2005; Balakrishnan et al. 2014).

In wild-type AAV 2, the expression of the replication proteins is regulated by the natural promoters p5 and p19. P5 drives the expression of Rep78 and Rep68, the latter being an alternative spliced version of the former. These two proteins possess specific endonuclease, ATPase and helicase activities; and they have vital functions in processes such as viral DNA replication, site-specific DNA integration, rescue of AAV DNA from the host genome and regulation of gene expression (Samulski et al. 2014; Im et al. 1992). The p19 promoter controls the expression Rep52 and Rep40, both proteins have helicase activity that is relevant for the accumulation and packaging of the DNA material (Smith et al. 1992; King et al. 2001).

Expression of the three VP capsid proteins is driven by the p40 promoter. Alternative splicing and the presence of conventional and non-conventional start codons influence the abundance of the resulting monomers. A fully assembled capsid is composed by 60 VP monomers at an approximate 1:1:10 molar ratio of VP1, VP2 and VP3, respectively. The N-terminal region of VP1 contains a unique phospholipase A2 motif necessary for viral infectivity (Popa-Wagner et al. 2012). These capsid proteins are beta barrel-folded, and are connected by long loops. The most prominent features of the capsid include a three-fold spike and a five-fold pore (Daya et al. 2008; Xie et al. 2002). Currently, thirteen AAV serotypes classified into eight groups have been defined based on their capsid amino acid sequence and serological patterns. In addition, the search for new variants in primate and non-primate tissues have led to the discovery of more than 100 variants. (Wu et al. 2006; Agbandje-McKenna et al. 2011).

Overview of wild-type AAV life cycle

As *Dependoparvovirus*, wild-type AAV requires cellular co-infection with a helper virus in order to achieve successful propagation inside a cell or tissue (Atchison et al. 1965). AAV has a defined manner to penetrate the cell and deliver the genetic material into the nucleus. Studies on AAV serotype 2 (the most studied and understood serotype) evidenced the affinity of AAV capsid to the heparin sulfate proteoglycan receptors present in different cell types (Summerford et al. 1998). The recognition process triggers endocytosis of the virus, which is capable of persisting the harsh endosomal environment. Changes in the endosomal internal pH modify the capsid structure, exposing the PLA2 motif and a nuclear localization signal (NLS). This latter event facilitates AAV endosome escape and nuclear entry (Stahnke et al. 2011; Barlett et al. 2000). Wild type AAV2 integrates its genome in a specific region of chromosome 19 called AAVS1. If no infection

with a helper virus occurs, AAV will undergo latency. On the contrary, helper virus infection triggers transcriptional activation of Rep proteins, which subsequently promotes excision of the provirus from the host genome and starts a cascade of events that lead to viral DNA replication and packaging of the viral genome. Finally, helper virus-mediated induced cell lysis allows release of newly formed AAV particles (Goncalves et al. 2005; Samulski et al 1991).

AAV DNA replication and encapsidation

Wild-type AAV DNA replication is based on a self-priming, strand displacement mechanism. Cellular polymerase recognizes the free 3 OH' end of one ITR and uses it as a primer for DNA elongation, forming duplex monomers. These duplex structures are then nicked by Rep78 at specific sequences within the ITR called terminal recognition sites (TRS) (Bristel et al. 2000). The terminal resolution process allows the unwinding of the ITRs and copying of these sequences into the daughter strand, generating two linear strands a result. This configuration promotes the renaturation of the ITR hairpin structure in both strands, generating incomplete duplex structures that will get polymerized via the same self-priming principle described previously. However, polymerization of the duplex triggers the displacement of the daughter strand, whereas the parental strand undergoes consecutive rounds of replication. In AAV and adenovirus coinfection experiments, it has been found that a group of cellular components, including DNA polymerase delta, minichromosome maintenance complex (MCM), replication factors A and C (RFA, RFC), and proliferating cell nuclear antigen (PCNA) are instrumental for the AAV DNA replication (Nash et al. 2007; Nash et al. 2008).

The exact mechanism on how single stranded AAV DNA gets encapsidated is not fully understood. However, some studies have provided an understanding of key steps during AAV DNA packaging. A “Rep-mediated encapsidation” model, suggests a strong interaction among pre-formed capsids, Rep proteins and single-stranded DNA (King et al. 2001; Samulski et al. 2014; Ling et al. 2015). Rep78/68 remains bound to the daughter strand after replication, and it is believed that this protein contribute to packaging by bringing the DNA close to the empty capsid. Then, it is hypothesized that Rep52/40 proteins interact with both capsid and DNA, and drive the insertion of DNA through the five-fold pore via their helicase activity. Studies have suggested an important role of the AAV ITR “D-sequence” in DNA packaging, nevertheless its function in this process is not clear (Wang et al. 1996).

Helper virus functions for successful AAV replication

Wild-type AAV propagation in cells requires super infection with a helper virus. Adenovirus, herpes simplex virus (HSV), vaccinia virus and human papillomavirus (HPV) have been reported to contribute to AAV proliferation in two main ways: By promoting the upregulation of AAV natural promoters that otherwise would be repressed, and by modifying the cellular milieu in such manner that host cell proteins do not interfere with AAV replication (Geoffroy et al. 2005; Daya et al. 2008). Several studies carried out on adenovirus and HSV helped to elucidate the specific set of genes that allows AAV lytic stage. The adenoviral set includes E1A, E2A, E1b, E4orf6 and VA RNA. Main activities of these genes include transcriptional activation of Rep genes, promotion of DNA replication, and transport and stability of AAV mRNA (Chang et al. 1990; Samulski et al. 1988). During AAV-Adenovirus coinfection, AAV uses the cellular DNA polymerase for the elongation of its DNA. The HSV helper set includes UL29, the main single-stranded DNA binding

protein, and the UL5/UL8/UL22 helicase-primase complex; these proteins contribute to AAV DNA recruitment into the replication centers. UL30/UL42 encodes for the HSV DNA polymerase which contributes to AAV DNA replication, but it has proven to be non-essential for the later purpose (Weindler et al. 1991; Slanina et al. 2006).

Molecular design for recombinant AAV vector production

Introduction

Recombinant adeno-associated viruses (rAAV) have gained increasing attention in the viral gene therapy field. A safe clinical profile, availability of viral serotypes with different tissue tropisms, and potential long-term gene expression are main advantages of rAAV as viral vector. To date, more than 100 gene therapy clinical trials have been conducted, tackling a variety of diseases such as Lipoprotein Lipase (LPL) Deficiency, Cystic Fibrosis, and hemophilic disorders (Carter 2005; Bryant et al. 2013; Gene therapy clinical trials website: http://www.abedia.com/wiley/)

Production of rAAV vectors started approximately 32 years ago, after different groups demonstrated formation of genetically modified AAV viral particles capable of infecting and transducing mammalian cells (Tratschin et al. 1984; Hermonat & Muczycska 1984). From a clinical perspective, these findings raised the possibility for developing a new type of therapeutic DNA vector. From a biotechnology and bioprocessing perspective, this new biologic represented a unique market opportunity, assuming that biotech manufacturing capabilities could meet quantity and quality requirements. Subsequently, significant effort has been put towards the design of simple and efficient processes for rAAV production.

Three expression systems are currently used for industrial vector production: Adenovirus (AdV), Herpesvirus (HSV) and Baculovirus (BV) complementation systems. Despite promising advances, there are several challenges associated with the manufacturing process. Complementation systems bring inherent complexity to the production process because rAAV formation requires an intricate interplay between virus and host genetic elements. This phenomenon has a direct effect on process robustness and understanding, as subtle variations on the number of biological or chemical inputs used in upstream bioprocessing can affect productivity. This problem is clearly seen in transfection-based protocols, where lot-to-lot yield can vary drastically based on the number and concentration of plasmids used, amount of transfection agent used, cell viability and mode of operation (Wright 2009; Huang et al. 2013; van der Loo and Wright 2016). Another concern during vector production refers to the generation of product-related impurities. Some components used during cell culture readily co-purify with rAAV or are difficult to remove without damaging the rAAV product, making purification challenging. Among them, collateral packaging of non-AAV DNA has raised concerns because of its potential clinical implications (Wright 2014). Finally, vector yield is one of the most limiting factors for potential commercial supply. Current specific productivity yields from different platforms range from 10^3 to 10^5 vector genomes (vg) per cell, and there is an ongoing effort to improve vector yields to satisfy high product demands (Clark 2002; Ayuso et al. 2010). Overall, the best molecular constructs are the ones where the nature of AAV genes, promoters and regulatory elements contribute to a simple, economic process that generates safe, high-quality vectors.

In this mini-review, the rational design of rAAV-producing expression systems is discussed, with special attention to molecular strategies that contribute to high-yielding, biomanufacturing-amenable rAAV production processes. In addition, in-depth details of a

new microbial system for rAAV vector production based on *Saccharomyces cerevisiae* are provided. The molecular configuration proposed by different groups and their potential implications in vector production processes are discussed. Other aspects of gene therapy vector bioprocessing such as cell line and inoculum scalability, transfection optimization, and media optimization are covered elsewhere (Negrete and Kotin 2008; Kotin 2011;Thomas et al. 2009;Clement and Grieger 2016; Robert et al. 2017).

rAAV vector biology

AAV is a non-enveloped, single-stranded DNA virus that belongs to the family *Parvoviridae*. The 4.7 kb genome contains three open reading frames, Rep, Cap and AAP (Fig 1a). Rep encodes four regulatory proteins (Rep78, Rep68, Rep52, and Rep40) that play important roles in replication and encapsidation of viral DNA. Cap encodes three capsid proteins (VP1, VP2, and VP3) and AAP encodes an assembly-activating protein that promotes capsid formation. The genome is flanked by inverted terminal repeats (ITRs) which contain Rep recognition sequences important for AAV DNA replication and packaging (Samulski and Muzycska 2014). As a *Dependoparvovirus*, AAV requires the aid of another virus to propagate in tissue culture. Adenovirus and Herpes Simplex virus have been traditionally used as AAV helper virus.

An AAV vector is a recombinant variant of the wild-type (wt) AAV virus, in which the natural coding and non-coding regions have been replaced by an expression cassette not bigger than 4.7kb. The vector genetic construct retains the lateral ITRs which are the only cis-acting elements required for replication and encapsidation of AAV DNA. With these modifications, rAAV becomes a replication-deficient entity, capable only of infecting cells and delivering DNA into their nuclei. As shown in fig 1b, vector production in culture

requires Rep and Cap genes to be provided in a separate construct (Kotterman and Schaffer 2014). AAV genetic elements can be re-arranged in multiple constructs, and then delivered into host cells via plasmid transfection or viral infection. Helper virus activities are also required for efficient vector production. They are provided by coinfection with helper virus stock, or transfection with a plasmid containing “helper” genes.

Adenovirus complementation system

Early AAV-producing systems contained three components: a plasmid with the transgene of interest flanked by ITRs; Rep and Cap genes expressed from wild type AAV or from a second plasmid; and AdV to provide helper functions (Tratschin et al. 1985). These systems showed positive AAV-like biological functionality (i.e AAV DNA replication, formation of AAV full particles, DNA rescue and replication after AdV coinfection) and served as a proof of concept for rAAV production. However, this process had limited potential for clinical use due to low yield and the presence of process-related impurities such as contaminating AdV and replication-competent (rc) AAV.

Subsequent approaches based exclusively on plasmid transfection removed the need for wild type AdV co-infection and usually require transfection with two or three plasmids containing the vector sequence (ITR-transgene-ITR), Rep/Cap genes, and helper virus auxiliary genes. Matsushita and collaborators (1998) evaluated different helper plasmid configurations in the search for the minimal set of genes required for AdV-free AAV vector production. Their best design included one plasmid with combined rAAV vector DNA sequence and Rep/Cap genes, and a second plasmid containing VA RNA, E4orf6, and E2A adenoviral genes. Because the HEK293 host cells already constitutively expressed the adenovirus proteins E1a and E1b, the full helper gene set included the 5

aforementioned proteins whose role on AAV production has been determined (Weitzman and Leiden 2011). This configuration tackled the initial problem of formation of replication-competent AAV and removed the need for infectious adenovirus. In parallel, reduction of rc-AAV formation was also accomplished by altering homologous sequences present in both vector and helper plasmids (Allen et al. 1997). Subsequent efforts aimed at reducing process and product-related impurities have been reported (Fig 2).

Despite design improvements, per cell productivity was lower than wild-type AAV viral yields (Clark 2002, Samulski & Muzyczka 1999). Several observations on natural AAV production indicated that the higher particle per cell yields might be linked to 1) Higher Rep and Cap gene copies per cell because of ITR-based DNA self-replication, and 2) Controlled expression of Rep78 levels have a positive impact on vector yield, and ameliorate Rep-mediated cytotoxicity which otherwise would impact cell viability (Schmidt et al 2000; Xiao et al 1998). Based on those premises, several groups proposed variations of the initial design. Xiao and collaborators (1998) used a two-plasmid system for enhanced vector production. The packaging plasmid, pXX2, contained an unconventional initiation codon ACG on the Rep gene to modulate its expression. In addition, this plasmid conformation included two p5 promoters (upstream of Rep and downstream of Cap) to improve p5' enhancer-like activity on the p40 promoter. The final yields obtained reached 10^5 particles per cell, which exceeded the values obtained from Matsushita by nearly 10-fold. Grimm and collaborators (1998) developed a pDG packaging plasmid, which combined Rep and Cap genes with AdV E2A, E4 and VA RNA helper genes. AAV p5 promoter was replaced by an MMTV-LTR promoter, which correlated to a reduction in Rep78 and increase in VP levels. Other molecular configurations were proposed by different groups, obtaining comparable vector yields (Allen et al. 2000; Collaco et al. 1999; Li and Samulski 2005; Table 1).

Such research efforts led to the adoption of the triple plasmid transfection system as one of the preferred methods for rAAV production. The standard configuration of this system provides the original Rep and Cap genes with their natural promoters (p5, p19, and p40), but striped from other cis-acting elements to reduce the probability of encapsidation. Under this configuration, one plasmid contains Rep and Cap genes, a second plasmid comprises the ITR-flanked transgene of interest, and a third plasmid provides helper genes (Fig 3a). Recently, Emmerling and collaborators (2016) evaluated a new Rep/Cap split packaging plasmid system, in which the original Rep and Cap genes were segregated into two plasmids. In the first plasmid, the Rep gene was split into two expression cassettes, one for Rep68 and another for Rep52 and Rep40. A second plasmid contained the Cap expression cassette for expression of the three VP proteins and AAP from the natural p40 promoter. This study indicated that high yield correlated to an optimal Rep68/Rep52 expression ratio and enhanced Cap expression. Overall, a 12-fold increase in vector productivity (vg/cell) and a 5-fold increase in transducing units/cell for this optimized plasmid transfection system was obtained, compared to the widely used pDG packaging system.

Although plasmid transfection methods offer simplicity and flexibility for basic research and early stage rAAV production, these protocols have limited scalability and reproducibility for use in large-scale production. Alternatively, stable packaging and producer cell lines were designed. These two different systems contain Rep/Cap only, or Rep/Cap plus vector constructs integrated into the cell genome, respectively. They all require the addition of Ad helper genes via plasmid transfection or virus infection, the latter being preferred to generate high yields (Chadeuf et al. 2000). This methodology improves cell culture process scalability and reduces the number of plasmids/virus required (Van der Loo and Wright 2016). Many variations of these stable cell lines have been designed

(Gao et al. 1998; Gao et al. 2002; Clark et al. 1995; Qiao et al. 2002) giving vector genome titers ranging around 10^4-10^6 vg/cell.

Herpesvirus complementation system

Early studies performed by Weindler and Heilbronn (1991) confirmed that HSV coinfection allowed AAV propagation, and that genes UL5/8/52 and DNA-binding protein gene UL29 were responsible for helper-like activities. The function of these HSV helper genes has also been determined (Weitzman and Leiden 2011). The HSV system became an alternative platform that could overcome production challenges identified in AdV-based systems, namely the complexity of large-scale transfection methods, as well as the presence of helper virus impurities. Early HSV plasmid designs used an HSV1 amplicon expressing AAV2 Rep and Cap proteins, plus wild-type HSV and a vector construct (Conway et al. 1997). This initial design was later improved by developing an ICP-27-deficient HSV strain which expressed AAV Rep/Cap (Conway et al. 1999). This recombinant strain was incapable of replicating in culture, reducing the generation of impurities throughout the process. Replication-deficient rHSV Rep/Cap propagation was performed as a separate process, by cultivating the virus in ICP27-expressing Vero cells. Efficient production of rAAV on HEK293 cells was accomplished by transducing cells with rHSV Rep/Cap and transfecting an AAV-GFP vector plasmid. Slightly better yields were observed when an HEK-derived, proviral cell line (GFP-92) was used. The use of replication-deficient rHSV shows no detectable levels of rc-AAV and low levels of viral helper particles in culture (Clement et al. 2009).

Current HSV-based design comprises two replication-deficient HSV strains engineered to individually harbor Rep/Cap and AAV vector sequences (Fig 3b). This transfection free-

approach was initially reported by Hwang and colleagues (2003), who were able to improve vector yield 30-fold relative to a transfection-based method after optimizing HSV multiplicity of infection (MOI) ratios (12 and 2 for rHSV/RepCap and rHSV/AAV-GFP viruses, respectively). Kang and collaborators (2009) expanded this rHSV-based, AAV production platform across multiple serotypes and transgenes with proven efficiency (higher than $1x10^5$ particles/cell). Other groups adapted the infection process to suspension culture. Per cell productivity obtained after infecting suspension-adapted BHK cells ranged around $8x10^4$ – $2x10^5$ vg/cell (Thomas et al. 2009; Knop et al. 2011). This modification allowed easier scale up and generated high yields. Current efforts are focused on improving viral inoculum growth and stability in culture.

Insect Cell - Baculovirus Expression System

The Baculovirus (BV) complementation system has become a reliable platform for expression of single heterologous proteins and multimeric particles. Using this method, recombinant viral strains usually derived from *Autographa Californica* Multinuclear Polyhedrosis Virus (AcMNPV) infect insect cells, hijacking the cellular machinery and expressing proteins encoded in the genome, including the foreign protein of interest. High per cell productivity yields are achieved because of BV's strong promoters. Insect cell lines such as Sf9 or Sf21 can grow in suspension and are easily adapted to scalable, stirred tank bioreactor-based processes (Merten et al. 2005). Evidence of human-like post translational modification capacity and intracellular viral capsid assembly support its use as a production platform for complex surface antigens, virus-like particles (VLPs), and potentially fully-assembled viral vectors (Van Oers et al. 2015; Fernandes et al. 2013).

Under these premises, several groups developed BV constructs that promoted AAV capsid formation, DNA replication and subsequent packaging within Sf9 cells. Urabe and collaborators (2002) designed an rAAV2 production strategy based on infection of Sf9 cells with three different recombinant BVs. The first BV carried Rep78 and Rep52 genes, the second carried the Cap gene, and the third carried the transgene flanked by ITRs. Preliminary experiments showed impaired Rep expression when AAV natural promoters were used; therefore, BV-specific promoters were adopted. Rep 78 and Rep52 expression were expressed from two independent cassettes controlled by baculovirus immediate early (ΔIE1) and polyhedrin (*polh*) promoters, respectively. The three VP proteins were expressed from a single cassette controlled by the *polh* promoter, and the VP1 start codon was mutated to ACG to enable expression of the three VP proteins at the appropriate ratio, without the need of alternative splicing. Per-cell productivity with this system achieved 5 x 10^4 vg/cell. The final product resembled HEK293-produced AAV vectors in its physical and biological properties. The same author later developed a BV infection process for production of rAAV5 that achieved considerable yield (~4-fold) and infectivity improvements by swapping in serotype 1 Rep52 for the serotype 5 analog, and the N-terminal portion of serotype 2 VP1 for the analogous portion from serotype 5 (Urabe et al. 2006). Chen (2008) further modified the system to allow expression of multiple proteins from single Rep and Cap coding sequences. *Polh* promoters were placed within intron sequences to drive expression of Rep and Cap short transcripts, allowing the production of Rep52, VP2, AAP and VP3 proteins. Rep78 and VP1 proteins were translated from longer transcripts after splicing of the internal *polh* promoter sequence. Under this concept, Rep78/52 and VP1/2/3 expression was accomplished by infection with either one or two BVs. Yields of 10^{14} vg/L were reached with the two-BV and three-BV proposed systems. Alternatively, Smith and collaborators (2009) were able to express Rep genes from a single mRNA species based on an mRNA leaky scanning mechanism, in which

several AUG codons found in the Rep 78 initial sequence (including the start codon) were mutated to suboptimal triplets, thus allowing translation of Rep52 by the ribosome. This concept was applied for both Rep and Cap, and the generated sequences were cloned in opposite transcriptional orientation and consolidated into one Rep/Cap BV.

Overall, the new two-BV design (Fig 3c) supported process robustness and scalable production of infectious rAAV by reducing the number of required BV viruses, and additionally increased virus stability. The latter aspect becomes relevant when using replication-competent baculovirus for vector production, as defective-interfering viral particles (DIPs) can emerge during culture. DIPs are variants that gain a competitive advantage against complete baculovirus by deleting subsets of genes. Outgrowth of DIPs correlates to loss of transgene and reduced rAAV productivity in insect cells. Strategies to prevent DIP formation during rAAV production include BV inoculum optimization at minimum multiplicity of infection (MOI), and molecular changes in the BV that reduce the number of homologous Rep sequences (Cecchini et al. 2008).

The generation of stable insect cell lines for production of recombinant AAV vectors has also been explored. Aslanidi and collaborators (2009) developed an inducible system that required infection with a single BV strain. The Rep and Cap sequences, controlled by baculovirus promoters were integrated within the Sf9 genome, and subsequent AAV gene amplification was triggered as a result of infection with a BV carrying an ITR-flanked transgene. The addition of a Rep Binding Element (RBE) upstream of Rep and Cap likely promoted feed-forward amplification of AAV gene expression. Overall vector yields were improved by 10-fold in comparison to the three-BV process. Mietzsch and collaborators (2014) further expanded Aslanidi's design and generated a production platform for rAAV serotypes 1-12 called OneBac. Rep genes from AAV2, 4, or 12 were used in combination with the Cap genes to generate yield improvements similar to Aslanidi (up to 5 x 10^5

vg/cell), but now for all serotypes. Mietzsch, however, noticed a decrease in VP1 levels and vector infectivity for rAAV5. Further improvements of this Sf9-infected stable production system (Mietzsch et al. 2015) based on the intron-splicing approach proposed by Chen (2008) led to a recovery of rAAV5 infectivity. Additionally, the authors were able to reduce collateral packaging of baculovirus genomic DNA by removing the previously integrated RBE signal.

The insect-cell BV expression system is one of the most promising platforms for recombinant AAV production. The Sf9 system's per cell productivity and volumetric productivity are among the highest documented to date ($>10^5$ vg/cell; ~10^{11} vg/mL) (Samulski and Muzyczka 2014). Efforts are currently underway to further improve cell specific productivity, BV stability, and process scalability.

AAV-producing yeast system

Saccharomyces cerevisiae is a unicellular, eukaryotic organism commonly used in research and technology. Baker's yeast's features include simple growth requirements, well-understood genetics, and post translational protein processing comparable to complex eukaryotic systems except subtle differences in N-linked glycosylation patterns (Nielsen 2013). These characteristics make *S. cerevisiae* suitable not only as a model system for eukaryotic cell biology studies, but also as a heterologous protein expression system for biotechnology applications. *Saccharomyces*' value has been explored in the pharmaceutical industry. Recombinant strains capable of producing therapeutic proteins, antigenic proteins, and virus-like particles (VLPs) have been generated. Its proven record of heterologous viral protein expression, evidence of assembled capsid production, and evidence of virus replication of some entities such as *Parvovirus* suggested the possibility

of generating full AAV viral particles in this organism (Kim and Kim 2016; Bill 2015; Zhao and Frazer 2002a; Zhao and Frazer 2002b).

Recent studies investigated rAAV generation in yeast. Backovic and collaborators (2012) utilized various plasmid constructs to demonstrate capsid protein expression and AAV capsid assembly in *Saccharomyces*. Their initial design placed the Rep and Cap genes under the control of their natural promoters along with an intron placed upstream of the VP1 initiation sequence. This design allowed the recovery of VP3 protein only. A second design with the Cap gene under a Gal1 promoter and a Kozak region upstream of VP1 facilitated expression of VP1. Co-transformation with the previously mentioned plasmids and optimized gene induction led to successful expression of VP1 and VP3 at a ratio comparable to other systems. Transmission electron microscopy studies confirmed capsid morphology of the purified product. Moreover, the same group demonstrated AAV single-stranded genome replication dependent on Rep expression and the presence of ITRs. DNA analysis of replicated sequences led the authors to suggest that AAV DNA replication process in yeast appear to be dissimilar from AAV canonical replication (Cervelli et al. 2011.)

Barajas and collaborators (2017) demonstrated production of fully-assembled, infectious AAV particles in *S. cerevisiae*. The system was based on four plasmids containing individual expression cassettes for two Rep proteins (Rep78 and Rep52), three VP capsids, and the assembly-activating protein (AAP) (Fig 3d). Unlike previous efforts, protein expression from the 6 expression cassettes was controlled by yeast-specific, galactose-inducible GAL promoters of different strengths, and codon optimization was required on Rep and AAP sequences. Southern blot analysis demonstrated formation of AAV DNA monomeric forms, and Western blot analysis of purified capsids showed detectable levels of the three VP capsids. Further examination confirmed transgene

presence and infectious capacity of the purified material. The authors reported rAAV2 full particle titer yield that ranges around 10^8vg/mL. This 4-plasmid system design aligns with other similarly proposed plasmid configurations (Takur 2002; Snyder 2011), and altogether, these studies constitute the proof-of-concept of rAAV production in a microbial system. The reported results not only demonstrate the potential utility of the yeast system as a tool to investigate AAV biology; this new concept raises the possibility of potentially developing an alternative, cost-effective, highly-scalable platform for rAAV production at large manufacturing scale. However, low vector yields and poor DNA encapsidation rates limit its application. Certain functionalities like those provided by helper viruses and host factors in other systems may be suboptimal in yeast and would need to be supplemented for high yield rAAV production. More investigations are needed to understand the benefits and limitations of this system that, although promising, is still far from becoming an efficient rAAV production platform.

Remarks

The molecular design of rAAV-producing expression systems plays a critical role not only in per cell vector productivity, but also in process robustness and elimination of process and product-related impurities. Each production system has advantages and disadvantages when adopted in pre-clinical or clinical manufacturing environments. AdV complementation systems based on triple plasmid transfection process are commonly adopted in lab settings because of their simple rAAV production workflow, flexibility to switch production to different AAV serotypes, and proven productivity and product quality. Scaling up this system, however, brings challenges associated to cell adaptation to suspension culture, plasmid generation and lot-lo-lot variability in transfection efficiency. The HSV complementation system based on the use of rHSVs overcomes scale-up

limitations regarding to cell culture scale-up and process variability, while maintaining high productivity at large scale. One of the main challenges of this system relies on the viral inoculum stability and propagation. Production of AAV vectors in the Sf9/BV complementation system has proven to be very efficacious, and several studies demonstrated the suitability of this system for large scale vector production. However, like other systems based on viral infection, special attention needs to be focused on the viral inoculum stability and generation. Additionally, more studies need to be performed to fully characterize quality attributes of Sf9-based vector material (Wang et al. 2011).

All rAAV expression systems have three fundamental commonalities: 1) successful delivery and amplification of the necessary genetic material into the host cell line; 2) fine-tuning of Rep-Cap expression levels; and 3) modification of the cellular milieu to a more "AAV-friendly" environment—something that has only recently begun to be investigated. The use of non-native hosts for rAAV production brings also the need to optimize critical variables such as timing and strength of expression of AAV components.

Besides molecular optimization of AAV genes, some research has focused on identifying limitations imposed by the host cells on rAAV production, or the negative effects caused by rAAV on the cells. Satkunanathan and collaborators (2014) identified Y-box binding protein (YB1) as an inhibitory protein of rAAV production. They postulated this protein could interfere with viral DNA encapsidation by competitive binding for recognition sequences. Experimental silencing of this gene translated into a significant increase in rAAV2 titer relative to control condition. Reid and colleagues (2017) hypothesized producer cells could experience cytotoxicity driven by rAAV transgene overexpression. mRNA silencing studies on different transgenes resulted in yield improvement up to 22-fold relative to control. Future research efforts in multiple directions will hopefully translate to highly productive vector expression systems.

S. cerevisiae as an experimental system for viral protein expression

Yeast as a model to understand viral replication

The use of model organisms has paramount importance in basic biology. They constitute simple, well-characterized organisms with conserved metabolic pathways that resemble biological phenotypes from more complex species. *Saccharomyces* is a popular eukaryotic model for cell biology research. Of its 6000+ genes, more than 70% have an identified function, and roughly half of the yeast proteome have human or plant orthologues (Nagy 2014). The use of a yeast model shortens the study time considerably because of yeast's high growth rates. Other features such as simple growth requirements and the availability of screening libraries tend to simplify the experimental design.

A handful of groups have used yeast as a model to understand virus-host interactions. These studies were focused on animal and plant viruses, and specifically cover the functional analysis of host cell factors in several steps of virus life cycle. The availability of many genetic and proteomic tools allowed investigation of some animal and plant RNA(+) virus in yeast models. Jiang (2006) utilized the yTHC collection, a screening library with inducible promoters upstream of selected essential genes, to identify host factors that impact Tombusvirus RNA replication. Thirty host genes whose downregulation either increase or decrease RNA replication were identified, many of them involved RNA transcription, protein transport, and other metabolic processes. Serviene (2006) used the same tool to study Tombusvirus RNA recombination. She identified 16 leads whose downregulation alter the ratio of recombinant/non-recombinant Tomato bunshy stunt virus (TBSV) RNA. Li (2008, 2009) used proteomic approaches to understand the role of host

RNA-binding proteins in virus RNA replication. Specifically, he determined the potential role of translation elongation factor 1a and Cdc34p as part of the replicase complex.

Yeast models have also been applied to study DNA viruses. Zhao and Fraser (2002a, 2002b) evidenced full replication and formation of infectious DNA virions after adding bovine papillomavirus (BPV) particles into yeast protoplasts. Angelleti (2002) developed a yeast model to study Human Papillomavirus (HPV) replication. By construction of different plasmids, the author was able to express viral proteins necessary for replication of multiple serotypes of HPV. Moreover, studies to depict the replication cycle of Geminiviruses in plants have been executed, being the indian mug bean yellow mosaic virus (IMBYMV) the most relevant virus from this group from the agricultural perspective. These studies include the use of yeast models to express viral proteins and evaluate the impact of host factors and external components on viral replication (Raghavan 2003; Chilakamarthi 2007; Kaliappan 2012).

S. cerevisiae as a heterologous host

S. cerevisiae strains are useful workhorses for production of recombinant products. In fermentation processes, this yeast shows fast growth rates, high scalability and simple media requirements. As an eukaryotic system, *Saccharomyces* is capable of performing post-translational modifications necessary to assure proper molecule folding and specific biological activities for the product of interest (Nielsen 2013). Over the last 30 years, academic and industry labs have developed recombinant strains of *S. cerevisiae* capable of producing a plethora of molecules, ranging from food enzymes to therapeutic proteins. One key to the success of *S.cerevisiae* as heterologous host relies on the variety of promoters and molecular elements that can be used advantageously for expression of

foreign proteins. The availability of several constitutive and inducible promoters, as well as molecular tools necessary for extra and intra-chromosomal gene expression, permits the engineering of multiple expression cassettes that can be modulated in strength and activation time (Darby 2012).

Another milestone that contributed to the usefulness of recombinant *S. cerevisiae* strains is the sequencing of its genome. The identification of more than 6000 genes in *S. cerevisiae* contributed to the understanding of several routes that compose yeast metabolism, and allowed scientists to target specific processes that directly or indirectly affect foreign protein production. Recent efforts to improve protein production focus on protein secretion, protein folding, stress, and cell cycle (Nielsen 2013).

Production of virus-like particles in S. cerevisiae

Virus-like particles (VLP) are polymeric, complex structures formed by one or many structural viral proteins. Once expressed, the monomers get assembled into a capsid structure that resembles a natural viral capsid. VLPs do not contain their respective viral nucleic acid, limiting their capacity to replicate or perform a biological activity proper of their particular wild-type virus' life cycle (Lua et al. 2014).

VLPs have multiple applications in biotechnology and the pharmaceutical industry. Because of its non-replicative capacity, these particles can be used as recombinant vaccines for animal or human use. Different molecular configurations of VLPs are used to exposed viral antigens to the body and elicit immune response. VLPs can also be used as nucleic acid delivery systems. The fact that assembled capsids can contain foreign DNA or RNA sequences is showing increasing interest in the gene therapy field. Encapsidation of a transgene is a complex process not fully understood which might

involve the participation of not only viral elements but also host cell proteins. Generally, a delivery vector is produced in an expression system by insertion of all viral genetic elements, with the requirement that non-structural proteins must be provided *in trans* and the foreign nucleic acid of interest is linked to specific recognition sequences (cis-acting elements) to allow encapsidation. Production of nucleic acid delivery vectors in yeast has been demonstrated. Legendre (2005) used *S. cerevisiae* as a system to produce a recombinant MS2 bacteriophage carrying growth hormone mRNA. This study became a proof of concept for potential RNA vaccine production in yeast. Under a similar concept, Sun et al. (2011) encapsulated HIV-Gag mRNA in the same bacteriophage. In order to evaluate the efficacy of the vaccine, he used the purified material to infect BALb/c mice and the serological analysis showed antigen-specific immune response. Barajas et al (2017) generated recombinant AAV2-GFP infectious particles using *S. cerevisiae* as experimental platform. The analytical tests of the purified material demonstrated that the generated particles were able to infect mammalian cells and promote GFP transduction.

References

1. Agbandje-McKenna M, Kleinschmidt J (2011) AAV capsid structure and cell interactions. Methods Mol Biol 807:47–92 . doi: 10.1007/978-1-61779-370-7_3
2. Allen JM, Debelak DJ, Reynolds TC, Miller AD (1997) Identification and elimination of replication-competent adeno-associated virus (AAV) that can arise by nonhomologous recombination during AAV vector production. J Virol 71:6816–6822.
3. Allen JM, Halbert CL, Miller AD (2000) Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6. Mol Ther 1:88–95. doi: 10.1006/mthe.1999.0010

4. Angeletti PC, Kim K, Fernandes FJ, Lambert PF (2002) Stable replication of papillomavirus genomes in Saccharomyces cerevisiae. J Virol 76:3350–3358
5. Aslanidi G, Lamb K, Zolotukhin S (2009) An inducible system for highly efficient production of recombinant adeno-associated virus (rAAV) vectors in insect Sf9 cells. Proc Natl Acad Sci USA 106:5059–5064. doi: 10.1073/pnas.0810614106
6. Atchison RW, Casto BC, Hammon WM (1965) ADENOVIRUS-ASSOCIATED DEFECTIVE VIRUS PARTICLES. Science 149:754–756
7. Ayuso E, Mingozzi F, Bosch F (2010) Production, purification and characterization of adeno-associated vectors. Curr Gene Ther 10:423–436.
8. Backovic A, Cervelli T, Salvetti A, Zentilin L, Giacca M, Galli A (2012) Capsid protein expression and adeno-associated virus like particles assembly in *Saccharomyces cerevisiae*. Microb Cell Fact 11:124. doi: 10.1186/1475-2859-11-124
9. Balakrishnan B, Jayandharan GR (2014) Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. Curr Gene Ther 14:86–100
10. Barajas D, Aponte-Ubillus JJ, Akeefe H, Cinek T, Peltier J, Gold D (2017) Generation of infectious recombinant Adeno-associated virus in Saccharomyces cerevisiae. PLoS ONE 12:e0173010 . doi: 10.1371/journal.pone.0173010
11. Bartlett JS, Wilcher R, Samulski RJ (2000) Infectious Entry Pathway of Adeno-Associated Virus and Adeno-Associated Virus Vectors. J Virol 74:2777–2785 . doi: 10.1128/JVI.74.6.2777-2785.2000
12. Bill RM (2015) Recombinant protein subunit vaccine synthesis in microbes: a role for yeast? J Pharm Pharmacol 67:319–328. doi: 10.1111/jphp.12353
13. Booth MJ, Mistry A, Li X, Thrasher A, Coffin RS (2004) Transfection-free and scalable recombinant AAV vector production using HSV/AAV hybrids. Gene Ther 11:829–837. doi: 10.1038/sj.gt.3302226

14. Bryant LM, Christopher DM, Giles AR, Hinderer C, Rodriguez JL, Smith JB, Traxler EA, Tycko J, Wojno AP, Wilson JM (2013) Lessons Learned from the Clinical Development and Market Authorization of Glybera. Hum Gene Ther Clin Dev 24:55–64. doi: 10.1089/humc.2013.087
15. Brister JR, Muzyczka N (2000) Mechanism of Rep-mediated adeno-associated virus origin nicking. J Virol 74:7762–7771
16. Carter BJ (2005) Adeno-Associated Virus Vectors in Clinical Trials. Hum Gene Ther 16:541–550. doi: 10.1089/hum.2005.16.541
17. Cecchini S, Negrete A, Kotin RM (2008) Toward exascale production of recombinant adeno-associated virus for gene transfer applications. Gene Ther 15:823–830. doi: 10.1038/gt.2008.61
18. Cervelli T, Backovic A, Galli A (2011) Formation of AAV Single Stranded DNA Genome from a Circular Plasmid in *Saccharomyces cerevisiae*. PLOS ONE 6:e23474. doi: 10.1371/journal.pone.0023474
19. 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. J Gene Med 2:260–268. doi: 10.1002/1521-2254(200007/08)2:4<260::AID-JGM111>3.0.CO;2-8
20. Chandler M, de la Cruz F, Dyda F, Hickman AB, Moncalian G, Ton-Hoang B (2013) Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. Nat Rev Microbiol 11:525–538. doi: 10.1038/nrmicro3067
21. Chang LS, Shenk T (1990) The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters. J Virol 64:2103–2109
22. Chen H (2008) Intron Splicing–mediated Expression of AAV Rep and Cap Genes

and Production of AAV Vectors in Insect Cells. Mol Ther 16:924–930. doi: 10.1038/mt.2008.35

23. Chilakamarthi U, Mukherjee SK, Deb JK (2007) Intervention of geminiviral replication in yeast by ribozyme mediated downregulation of its Rep protein. FEBS Letters 581:2675–2683 . doi: 10.1016/j.febslet.2007.04.084
24. Clark KR (2002) Recent advances in recombinant adeno-associated virus vector production. Kidney Int 61:S9-15. doi: 10.1046/j.1523-1755.2002.0610s1009.x
25. Clark KR, Voulgaropoulou F, Fraley DM, Johnson PR (1995) Cell lines for the production of recombinant adeno-associated virus. Hum Gene Ther 6:1329–1341. doi: 10.1089/hum.1995.6.10-1329
26. Clément N, Grieger JC (2016) Manufacturing of recombinant adeno-associated viral vectors for clinical trials. Mol Ther Methods Clin Dev 3:16002. doi: 10.1038/mtm.2016.2
27. Clément N, Knop DR, Byrne BJ (2009) Large-scale adeno-associated viral vector production using a herpesvirus-based system enables manufacturing for clinical studies. Hum Gene Ther 20:796–806. doi: 10.1089/hum.2009.094
28. Collaco RF, Cao X, Trempe JP (1999) A helper virus-free packaging system for recombinant adeno-associated virus vectors. Gene 238:397–405.
29. Conway JE, Rhys CM, Zolotukhin I, Zolotukhin S, Muzyczka N, Hayward GS, Byrne BJ (1999) High-titer recombinant adeno-associated virus production utilizing a recombinant herpes simplex virus type I vector expressing AAV-2 Rep and Cap. Gene Ther 6:986–993. doi: 10.1038/sj.gt.3300937
30. Conway JE, Zolotukhin S, Muzyczka N, Hayward GS, Byrne BJ (1997) Recombinant adeno-associated virus type 2 replication and packaging is entirely supported by a herpes simplex virus type 1 amplicon expressing Rep and Cap. J Virol 71:8780–8789.

31. Darby RAJ, Cartwright SP, Dilworth MV, Bill RM (2012) Which yeast species shall I choose? Saccharomyces cerevisiae versus Pichia pastoris (review). Methods Mol Biol 866:11–23 . doi: 10.1007/978-1-61779-770-5_2
32. Daya S, Berns KI (2008) Gene Therapy Using Adeno-Associated Virus Vectors. Clin Microbiol Rev 21:583–593 . doi: 10.1128/CMR.00008-08
33. Dong B, Moore AR, Dai J, Roberts S, Chu K, Kapranov P, Moss B, Xiao W (2013) A concept of eliminating nonhomologous recombination for scalable and safe AAV vector generation for human gene therapy. Nucleic Acids Res 41:6609–6617. doi: 10.1093/nar/gkt404
34. Emmerling VV, Pegel A, Milian EG, Venereo-Sanchez A, Kunz M, Wegele J, Kamen AA, Kochanek S, Hoerer M (2016) Rational plasmid design and bioprocess optimization to enhance recombinant adeno-associated virus (AAV) productivity in mammalian cells. Biotechnol J 11:290–297. doi: 10.1002/biot.201500176
35. Fernandes F, Teixeira AP, Carinhas N, Carrondo MJ, Alves PM (2013) Insect cells as a production platform of complex virus-like particles. Expert Rev Vaccines 12:225–236. doi: 10.1586/erv.12.153
36. Gao G, Lu F, Sanmiguel JC, Tran PT, Abbas Z, Lynd KS, Marsh J, Spinner NB, Wilson JM (2002) Rep/Cap gene amplification and high-yield production of AAV in an A549 cell line expressing Rep/Cap. Mol Ther 5:644–649. doi: 10.1006/mthe.2001.0591
37. Gao GP, Qu G, Faust LZ, Engdahl RK, Xiao W, Hughes JV, Zoltick PW, Wilson JM (1998) High-titer adeno-associated viral vectors from a Rep/Cap cell line and hybrid shuttle virus. Hum Gene Ther 9:2353–2362. doi: 10.1089/hum.1998.9.16-2353
38. Geoffroy M-C, Salvetti A (2005) Helper functions required for wild type and recombinant adeno-associated virus growth. Curr Gene Ther 5:265–271

39. Gonçalves MA (2005) Adeno-associated virus: from defective virus to effective vector. Virology Journal 2:43 . doi: 10.1186/1743-422X-2-43
40. Grimm D, Kern A, Rittner K, Kleinschmidt JA (1998) Novel Tools for Production and Purification of Recombinant Adeno associated Virus Vectors. Hum Gene Ther 9:2745–2760. doi: 10.1089/hum.1998.9.18-2745
41. Hauck B, Murphy SL, Smith PH, Qu G, Liu X, Zelenaia O, Mingozzi F, Sommer JM, High KA, Wright JF (2009) Undetectable transcription of cap in a clinical AAV vector: implications for preformed capsid in immune responses. Mol Ther 17:144–152. doi: 10.1038/mt.2008.227
42. Hermonat PL, Muzyczka N (1984) Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. Proc Natl Acad Sci USA 81:6466–6470.
43. Huang X, Hartley A-V, Yin Y, Herskowitz JH, Lah JJ, Ressler KJ (2013) AAV2 production with optimized N/P ratio and PEI-mediated transfection results in low toxicity and high titer for in vitro and in vivo applications. J Virol Methods 193:270–277. doi: 10.1016/j.jviromet.2013.06.008
44. Hwang K, Mandell T, Kintner H, Zolotukhin S, Snyder RO, Byrne BJ (2003) High Titer Recombinant Adeno-Associated Virus Production Using Replication Deficient Herpes Simplex Viruses Type I. Mol Ther 7:S14–S15. doi: 10.1016/S1525-0016(16)40477-6
45. Im DS, Muzyczka N (1992) Partial purification of adeno-associated virus Rep78, Rep52, and Rep40 and their biochemical characterization. J Virol 66:1119–1128
46. Jiang Y, Serviene E, Gal J, Panavas T, Nagy PD (2006) Identification of Essential Host Factors Affecting Tombusvirus RNA Replication Based on the Yeast Tet Promoters Hughes Collection. J Virol 80:7394–7404 . doi: 10.1128/JVI.02686-05

47. Kaliappan K, Choudhury NR, Suyal G, Mukherjee SK (2012) A novel role for RAD54: this host protein modulates geminiviral DNA replication. FASEB J 26:1142–1160 . doi: 10.1096/fj.11-188508
48. Kang W, Wang L, Harrell H, Liu J, Thomas DL, Mayfield TL, Scotti MM, Ye GJ, Veres G, Knop DR (2009) An efficient rHSV-based complementation system for the production of multiple rAAV vector serotypes. Gene Ther 16:229–239. doi: 10.1038/gt.2008.158
49. Kim HJ, Kim H-J (2016) Yeast as an expression system for producing virus-like particles: what factors do we need to consider? Lett Appl Microbiol 64:111–123. doi: 10.1111/lam.12695
50. King JA, Dubielzig R, Grimm D, Kleinschmidt JA (2001) DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. EMBO J 20:3282–3291 . doi: 10.1093/emboj/20.12.3282
51. Knop DR, Thomas DL, Butts, C (2011) SFM rHSV-Based Production of rAAV Vectors in Suspension BHK Cells. Mol Ther 19:S227. doi: 10.1016/S1525-0016(16)37166-0
52. Kotin RM (2011) Large-scale recombinant adeno-associated virus production. Hu Mol Genet 20:R2–R6. doi: 10.1093/hmg/ddr141
53. Kotterman MA, Schaffer DV (2014) Engineering adeno-associated viruses for clinical gene therapy. Nat Rev Genet 15:445–451. doi: 10.1038/nrg3742
54. Legendre D, Fastrez J (2005) Production in Saccharomycescerevisiae of MS2 virus-like particles packaging functional heterologous mRNAs. J Biotechnol 117:183–194 . doi: 10.1016/j.jbiotec.2005.01.010
55. Li Z, Barajas D, Panavas T, Herbst DA, Nagy PD (2008) Cdc34p ubiquitin-conjugating enzyme is a component of the tombusvirus replicase complex and

ubiquitinates p33 replication protein. J Virol 82:6911–6926 . doi: 10.1128/JVI.00702-08

56. Li C, Samulski RJ (2005) Serotype-specific replicating AAV helper constructs increase recombinant AAV type 2 vector production. Virology 335:10–21. doi: 10.1016/j.virol.2005.02.008
57. Li Z, Pogany J, Panavas T, Xu K, Esposito AM, Kinzy TG, Nagy PD (2009) Translation elongation factor 1A is a component of the tombusvirus replicase complex and affects the stability of the p33 replication co-factor. Virology 385:245–260 . doi: 10.1016/j.virol.2008.11.041
58. Ling C, Wang Y, Lu Y, Wang L, Jayandharan GR, Aslanidi GV, Li B, Cheng B, Ma W, Lentz T, Ling C, Xiao X, Samulski RJ, Muzyczka N, Srivastava A (2015) The Adeno-Associated Virus Genome Packaging Puzzle. J Mol Genet Med 9: . doi: 10.4172/1747-0862.1000175
59. Lua LHL, Connors NK, Sainsbury F, Chuan YP, Wibowo N, Middelberg APJ (2014) Bioengineering virus-like particles as vaccines. Biotechnol Bioeng 111:425–440 . doi: 10.1002/bit.25159
60. Matsushita T, Elliger S, Elliger C, Podsakoff G, Villarreal L, Kurtzman GJ, Iwaki Y, Colosi P (1998) Adeno-associated virus vectors can be efficiently produced without helper virus. Gene Ther 5:938–945. doi: 10.1038/sj.gt.3300680
61. Merten O-W, Gény-Fiamma C, Douar AM (2005) Current issues in adeno-associated viral vector production. Gene Ther 12:S51–S61. doi: 10.1038/sj.gt.3302615
62. Mietzsch M, Casteleyn V, Weger S, Zolotukhin S, Heilbronn R (2015) OneBac 2.0: Sf9 Cell Lines for Production of AAV5 Vectors with Enhanced Infectivity and Minimal Encapsidation of Foreign DNA. Hum Gene Ther 26:688–697. doi: 10.1089/hum.2015.050

63. Mietzsch M, Grasse S, Zurawski C, Weger S, Bennett A, Agbandje-McKenna M, Muzyczka N, Zolotukhin S, Heilbronn R (2014) OneBac: platform for scalable and high-titer production of adeno-associated virus serotype 1-12 vectors for gene therapy. Hum Gene Ther 25:212–222. doi: 10.1089/hum.2013.184
64. Nagy PD, Pogany J, Lin J-Y (2014) How yeast can be used as a genetic platform to explore virus-host interactions: from “omics” to functional studies. Trends Microbiol 22:309–316 . doi: 10.1016/j.tim.2014.02.003
65. Nash K, Chen W, McDonald WF, Zhou X, Muzyczka N (2007) Purification of host cell enzymes involved in adeno-associated virus DNA replication. J Virol 81:5777–5787 . doi: 10.1128/JVI.02651-06
66. Nash K, Chen W, Muzyczka N (2008) Complete In Vitro Reconstitution of Adeno-Associated Virus DNA Replication Requires the Minichromosome Maintenance Complex Proteins. J Virol 82:1458–1464 . doi: 10.1128/JVI.01968-07
67. Negrete A, Kotin RM (2008) Strategies for manufacturing recombinant adeno-associated virus vectors for gene therapy applications exploiting baculovirus technology. Brief Funct Genomic Proteomic 7:303–311. doi: 10.1093/bfgp/eln034
68. Nielsen J (2013) Production of biopharmaceutical proteins by yeast. Bioengineered 4:207–211 . doi: 10.4161/bioe.22856
69. Popa-Wagner R, Porwal M, Kann M, Reuss M, Weimer M, Florin L, Kleinschmidt JA (2012) Impact of VP1-Specific Protein Sequence Motifs on Adeno-Associated Virus Type 2 Intracellular Trafficking and Nuclear Entry. J Virol 86:9163–9174 . doi: 10.1128/JVI.00282-12
70. 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. J Virol 76:13015–13027. doi: 10.1128/JVI.76.24.13015-13027.2002

71. Raghavan V, Malik PS, Choudhury NR, Mukherjee SK (2004) The DNA-A Component of a Plant Geminivirus (Indian Mung Bean Yellow Mosaic Virus) Replicates in Budding Yeast Cells. J Virol 78:2405–2413 . doi: 10.1128/JVI.78.5.2405-2413.2004
72. Reid CA, Boye SL, Hauswirth WW, Lipinski DM (2017) miRNA-mediated post-transcriptional silencing of transgenes leads to increased adeno-associated viral vector yield and targeting specificity. Gene Ther. doi: 10.1038/gt.2017.50
73. Robert M-A, Chahal PS, Audy A, Kamen A, Gilbert R, Gaillet B (2017) Manufacturing of recombinant adeno-associated viruses using mammalian expression platforms. Biotechnol J 12:1600193. doi: 10.1002/biot.201600193
74. Samulski RJ, Muzyczka N (2014) AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. Annual Review of Virology 1:427–451. doi: 10.1146/annurev-virology-031413-085355
75. Samulski RJ, Shenk T (1988) Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. J Virol 62:206–210
76. Samulski RJ, Zhu X, Xiao X, Brook JD, Housman DE, Epstein N, Hunter LA (1991) Targeted integration of adeno-associated virus (AAV) into human chromosome 19. EMBO J 10:3941–3950
77. Samulski RJ, Muzyczka N (1999) Adeno-associated Viral Vectors. In Friedmann T (ed) The Development of Gene Therapy. Cold Spring Harbor, New York, pp 131-172.
78. Satkunanathan S, Wheeler J, Thorpe R, Zhao Y (2014) Establishment of a Novel Cell Line for the Enhanced Production of Recombinant Adeno-Associated Virus Vectors for Gene Therapy. Hum Gene Ther 25:929–941. doi: 10.1089/hum.2014.041
79. Schmidt M, Afione S, Kotin RM (2000) Adeno-Associated Virus Type 2 Rep78

Induces Apoptosis through Caspase Activation Independently of p53. J Virol 74:9441–9450

80. Serviene E, Jiang Y, Cheng C-P, Baker J, Nagy PD (2006) Screening of the yeast yTHC collection identifies essential host factors affecting tombusvirus RNA recombination. J Virol 80:1231–1241 . doi: 10.1128/JVI.80.3.1231-1241.2006
81. Slanina H, Weger S, Stow ND, Kuhrs A, Heilbronn R (2006) Role of the Herpes Simplex Virus Helicase-Primase Complex during Adeno-Associated Virus DNA Replication. J Virol 80:5241–5250 . doi: 10.1128/JVI.02718-05
82. Smith RH, Kotin RM (1998) The Rep52 gene product of adeno-associated virus is a DNA helicase with 3'-to-5' polarity. J Virol 72:4874–4881
83. Smith RH, Levy JR, Kotin RM (2009) A simplified baculovirus-AAV expression vector system coupled with one-step affinity purification yields high-titer rAAV stocks from insect cells. Mol Ther 17:1888–1896. doi: 10.1038/mt.2009.128
84. Snyder RO (2011) Production of recombinant AAV virions. US Patent 7,927,585 B2.
85. Stahnke S, Lux K, Uhrig S, Kreppel F, Hösel M, Coutelle O, Ogris M, Hallek M, Büning H (2011) Intrinsic phospholipase A2 activity of adeno-associated virus is involved in endosomal escape of incoming particles. Virology 409:77–83 . doi: 10.1016/j.virol.2010.09.025
86. Summerford C, Samulski RJ (1998) Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. J Virol 72:1438–1445
87. Sun S, Li W, Sun Y, Pan Y, Li J (2011) A new RNA vaccine platform based on MS2 virus-like particles produced in Saccharomyces cerevisiae. Biochem Biophys Res Commun 407:124–128 . doi: 10.1016/j.bbrc.2011.02.122
88. Thakur SS (2002) Production of Recombinant Adeno-Associated Viral Vectors in

Yeast. Dissertation. University of Florida. http://ufdc.ufl.edu/UFE0044702/00001

89. Thomas DL, Wang L, Niamke J, Liu J, Kang W, Scotti MM, Ye G-J, Veres G, Knop DR (2009) Scalable Recombinant Adeno-Associated Virus Production Using Recombinant Herpes Simplex Virus Type 1 Coinfection of Suspension-Adapted Mammalian Cells. Hum Gene Ther 20:861–870. doi: 10.1089/hum.2009.004
90. Tratschin JD, Miller IL, Smith MG, Carter BJ (1985) Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. Mol Cell Biol 5:3251–3260.
91. Tratschin JD, West MH, Sandbank T, Carter BJ (1984) A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. Mol Cell Biol 4:2072–2081.
92. Urabe M, Ding C, Kotin RM (2002) Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. Hum Gene Ther 13:1935–1943. doi: 10.1089/10430340260355347
93. Urabe M, Nakakura T, Xin K-Q, Obara Y, Mizukami H, Kume A, Kotin RM, Ozawa K (2006) Scalable generation of high-titer recombinant adeno-associated virus type 5 in insect cells. J Virol 80:1874–1885. doi: 10.1128/JVI.80.4.1874-1885.2006
94. van der Loo JCM, Wright JF (2016) Progress and challenges in viral vector manufacturing. Hum Mol Genet 25:R42-52. doi: 10.1093/hmg/ddv451
95. van Oers MM, Pijlman GP, Vlak JM (2015) Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. J Gen Virol 96:6–23. doi: 10.1099/vir.0.067108-0
96. Wang XS, Ponnazhagan S, Srivastava A (1996) Rescue and replication of adeno-associated virus type 2 as well as vector DNA sequences from recombinant

plasmids containing deletions in the viral inverted terminal repeats: selective encapsidation of viral genomes in progeny virions. J Virol 70:1668–1677

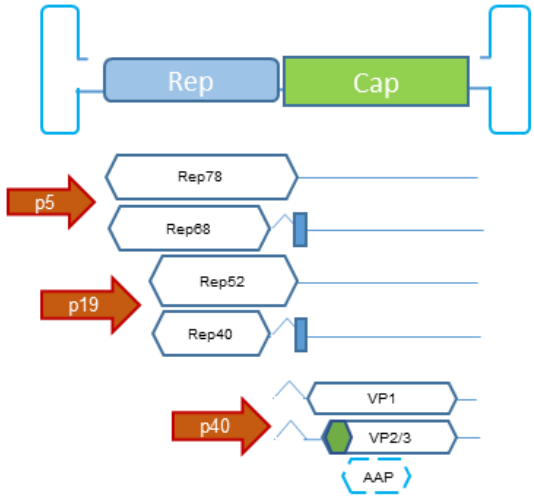
97. Wang L, Blouin V, Brument N, Bello-Roufai M, Francois A (2011) Production and purification of recombinant adeno-associated vectors. Methods Mol Biol 807:361–404 . doi: 10.1007/978-1-61779-370-7_16
98. Weindler FW, Heilbronn R (1991) A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. J Virol 65:2476–2483
99. Weitzman MD, Linden RM (2011) Adeno-associated virus biology. Methods Mol Biol 807:1–23 . doi: 10.1007/978-1-61779-370-7_1
100. Wright JF (2009) Transient Transfection Methods for Clinical Adeno-Associated Viral Vector Production. Hum Gene Ther 20:698–706. doi: 10.1089/hum.2009.064
101. Wright JF (2014) Product-Related Impurities in Clinical-Grade Recombinant AAV Vectors: Characterization and Risk Assessment. Biomedicines 2:80–97. doi: 10.3390/biomedicines2010080
102. Wu Z, Asokan A, Samulski R (2006) Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy. Molecular Therapy 14:316–327. doi: 10.1016/j.ymthe.2006.05.009
103. Xiao X, Li J, Samulski RJ (1998) Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. J Virol 72:2224–2232.
104. Xie Q, Bu W, Bhatia S, Hare J, Somasundaram T, Azzi A, Chapman MS (2002) The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. PNAS 99:10405–10410 . doi: 10.1073/pnas.162250899

105. Zhao K-N, Frazer IH (2002a) Replication of bovine papillomavirus type 1 (BPV-1) DNA in Saccharomyces cerevisiae following infection with BPV-1 virions. J Virol 76:3359–3364

106. Zhao K-N, Frazer IH (2002b) Saccharomyces cerevisiae Is Permissive for Replication of Bovine Papillomavirus Type 1. J Virol 76:12265–12273 . doi: 10.1128/JVI.76.23.12265-12273.2002

Figures

a.



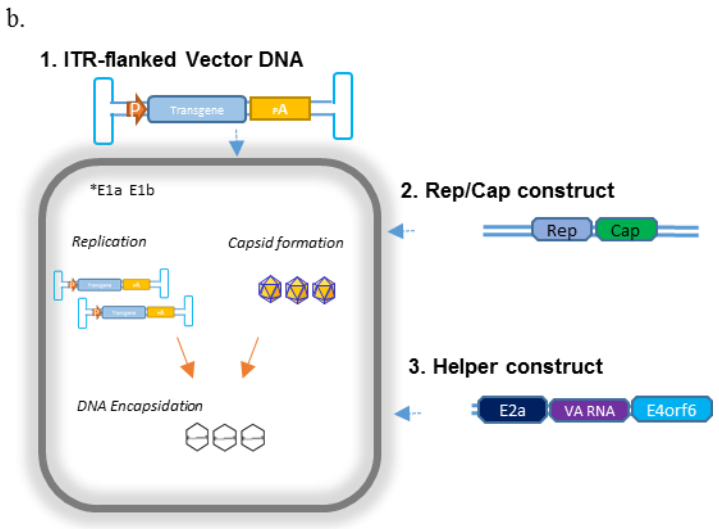


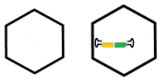
Fig 1. Adeno-associated virus (AAV) vector biology. Wild-type AAV genome (a) contains Rep and Cap genes. Rep encodes four regulatory proteins that play important roles in replication and encapsidation of viral DNA, and their expression is controlled by p5 and p19 promoters. Cap encodes three capsid proteins and assembly-activating protein (AAP), regulated by p40 promoter.

In an AAV vector (b), the wild-type AAV Rep and Cap genes have been replaced with the transgene of interest. Three components have to be delivered into the host cell line either by transfection or viral infection: Vector AAV DNA containing the transgene of interest, Rep and Cap genes (also known as packaging construct), and helper genes from Adenovirus. Rep78 and 68 promote AAV DNA rescue and subsequent replication. Cap proteins are synthesized in the cytoplasm and are shuttled to the nucleus for assembly. AAP supports assembly and maturation of the AAV capsid (Samulski & Muzyczka 2014). Rep52 and 40 interact with single-stranded DNA and pre-formed capsids to promote viral DNA encapsidation by a mechanism not yet fully understood (Ling et al. 2015). P: promoter; pA: polyadenylation sequence.

Process-related impurities

-Host and viral DNA
-Host and viral proteins
-Helper virus (*i.e.* AdV,HSV, BV)
-Replication-competent AAV
-Defective interfering viral particles (DIPs)

Product-related impurities



-Empty AAV capsids
-Full AAV capsids with foreign DNA sequences

Molecular strategies to minimize the presence of impurities
-Development of helper virus plasmids to minimize the presence of AdV particles in harvested supernatant (Matsushita et al. 1998; Grimm et al. 1998)
-Remove homologous sequences present between AAV vector and helper plasmids to prevent rc-AAV formation (Allen et al. 1997)
-Placement of Rep and Cap genes in opposite transcriptional orientation to prevent nonhomologous recombination-based rc-AAV formation (Allen et al. 1997)
-Use of an AAV vector oversized plasmid backbone (~6.9kb) to reduce the percentage of plasmid DNA insertion into formed AAV capsids (Hauck et al. 2008)
-Using a vaccinia virus-based carrier to deliver Rep and Cap genes in cytoplasm, reducing likelihood for homologous recombination events (Dong et al. 2013)
-Expression of AAV proteins from single Rep and Cap sequences prevented accumulation of DIPs, by reducing the amount of AAV Rep homologous sequences (Chen 2008)

Fig 2. rAAV production-related impurities and molecular strategies aimed for their reduction. Adapted from Wright (2014).

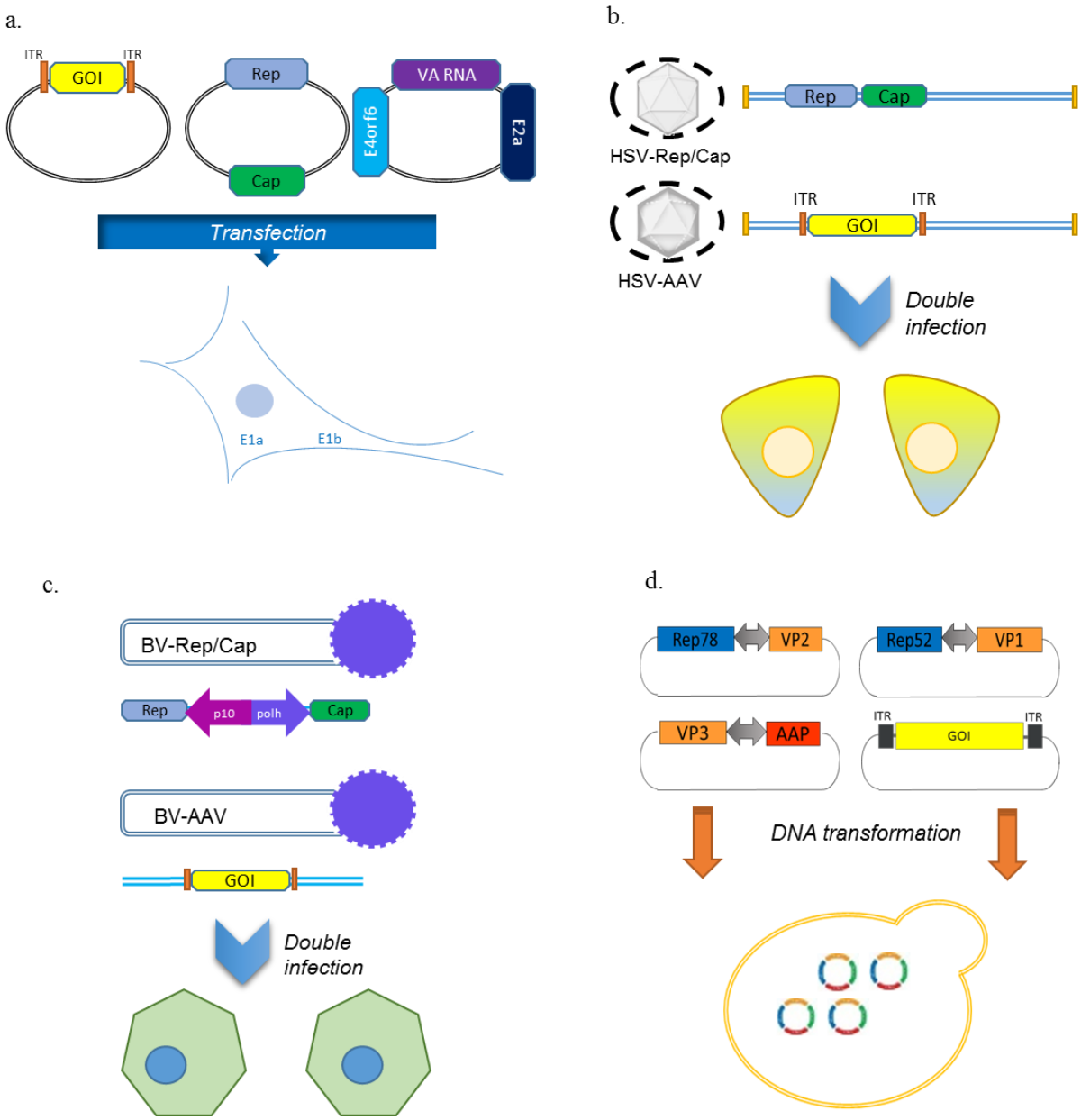


Fig 3. Diagram of rAAV-producing systems: Production in adenovirus complementation systems (a) are traditionally performed as plasmid transfection processes, where AAV Rep/Cap genes, the ITR-flanked gene of interest (GOI), as well as AdV-helper genes are provided to a E1a/E1b-containing HEK293 cell line. HSV complementation systems (b) use two recombinant herpes viral strains to provide AAV Rep/Cap genes, GOI, and HSV-helper elements to a mammalian cell line such as BHK. Sf9 - Baculovirus expression

systems (c) require two recombinant BV viral strains to provide the AAV-producing capability to insect cells. AAV protein expression is controlled by baculovirus promoters. Yeast-based systems (d) are transformed with a set of extrachromosomal plasmids that contain six AAV expression cassettes and GOI. AAV protein expression is controlled by yeast natural promoters.

Tables

Table 1. Reported rAAV vector yields on the articles cited in this chapter.

Method/Design	Vector yield	Reference
AdV-based system		
Ad-free, triple plasmid transfection	120ETU/cell	Matsushita et al. (1998)
Transient transfection, pXX2, unconventional start codon to modulate Rep expression	$1.2x10^3$ETU/cell-$9.4x10^5$vg/cell	Xiao et al. (1998)
Rep/Cap Hela stable cell line + AdV	up to 36IP/cell	Clark et al. (1995)
Transient transfection, pSH3/pSH5 plasmids which combined AAV rep,cap, and AdV-helper genes	$1.3x10^4$vg/cell 135ETU/cell	Collaco et al. (1999)
Transient transfection with pDG plasmid, MMLV regulate Rep expression	150IP/cell	Grimm et al. (1998)
Transient transfection, Mtrep-CMVcap plasmid, E4orf6-only plasmid	10^4vg/cell 23TU/cell	Allen et al. (2000)
Rep/Cap stable cell line A549 + AdV	262TU/cell	Gao et al. (2002)
Rep/Cap stable cell line + AdV-Cre vector	$1.3x10^5$vg/cell $1.7x10^3$ETU/cell[a]	Qiao et al. (2002)
Self-replicating Rep/Cap helper construct	$2x10^9$IU/well	Li and Samulski (2005)
Transient transfection, Rep/Cap split system	$2.6x10^5$vg/cell 37.8 TU/cell	Emmerling et al. (2016)

HSV-based system

Transient transfection, HSV-rc/d27 amplicon system	480vg/cell	Conway et al. (1997)
Infection with rHSV-rc strain on AAV-GFP-integrated cell line	480ETU/cell	Conway et al. (1999)
Infection with rHSV Rep/Cap and rHSV-GFP	$1.5x10^5$vg/cell $6x10^3$IP/cell	Hwang et al. (2003)
Double infection with ICP27-deleted rHSV strains	40TU/cell	Booth et al. (2004)
Production of rAAV serotypes 1,2, and 9, by double infection with ICP27-deleted rHSV strains	>$1.3x10^5$vg/cell >$9x10^3$IP/cell	Kang et al. (2009)
Production of rAAV serotypes 1,2,5 and 8, by using suspension-adapted BHK cells infected with rHSV strains	up to $1x10^5$vg/cell up to $1x10^4$IP/cell	Thomas et al. (2009)
Suspension-adapted BHK cells infected with rHSV strains	>$5x10^4$vg/cell	Knop et al. (2011)

BV-based system

Initial design, triple-BV system, Rep/Cap genes controlled by IE1/*polh* promoters	$5x10^4$vg/cell 30TU/cell	Urabe et al. (2002)
Swapping of AAV genetic elements from different serotypes to improve production of rAAV5	$6x10^4$vg/cell	Urabe et al. (2006)
Intron-splicing mediated expression	$1x10^{11}$vg/mL	Chen (2008)
Stable cell line, integration of Rep/Cap sequences + hr2-0.9 homologous regions + RBE sites	>10^5vg/cell	Aslanidi et al. (2009)
Stable cell line producing rAAV serotypes 1-12	up to $5x10^5$vg/cell	Mietzsch et al. (2014)

yeast expression system

4-plasmid system, individual expression cassettes regulated by Gal1/Gal10 promoters	~10^8 vg/ml	Barajas et al. (2017)

vg: vector genomes. ETU: Enhanced transducing unit. TU: Transducing unit. IP: Infectious particle. IU: infectious unit.

[a] Assuming 5 x 10^6 cells per 10 cm plate were used for assay

CHAPTER 3. RESEARCH OBJECTIVES

Main hypothesis:

"*Within the AAV-producing yeast platform, there are host cell factors and cell culture parameters that are impacting the efficiency of rAAV2 production. A thorough molecular and process design for this platform could potentially improve volumetric yield and specific productivity."*

Research Objectives:

1. Identify yeast host cell factors that exert an inhibitory or enhancing effect on rAAV2-GFP vector production.

2. Profile yeast proteome changes associated to rAAV2 vector production and identify potential links between these events and vector production constraints.

3. Design a fed-batch fermentation process for the production of AAV2-GFP.

CHAPTER 4.

Use of a rAAV2-producing yeast screening model to identify host proteins enhancing rAAV DNA replication and vector yield

ABSTRACT.

Background: Recombinant adeno-associated viral vectors (rAAV) are promising therapies for genetic diseases. Although current platforms for recombinant vector production are capable of generating drug material for pre-clinical and clinical studies, rAAV biomanufacturing will eventually face commercial supply challenges if per cell vector productivity and process scalability are not improved. Because considerable efforts have traditionally focused on optimizing rAAV plasmid design, herein we investigate the impact of host cell proteins on vector production to identify proteins that may enhance rAAV yield.

Results: Using a rAAV2-GFP-producing *Saccharomyces cerevisiae* model in combination with the yTHC screening library, we identified 22 gene candidates that improved rAAV DNA replication (rAAV-GFP/18s rDNA ratio) and vector yield (benzonase-resistant rAAV DNA vector genome titer) as high as 6-fold and 15-fold relative to control, respectively. The candidate proteins participate in biological processes such as DNA replication, ribosome biogenesis, and RNA and protein processing. The best 5 candidates (*PRE4, HEM4, TOP2, GPN3, SDO1*) were further screened by generating overexpression mutants in the YPH500 yeast strain. Subsequent clone evaluation was performed to confirm the rAAV-promoting activity of selected candidates under plate-based and bioreactor-controlled fermentation conditions. Digital droplet PCR analysis of cell lysate and AVB resin-purified material confirmed HEM4 and TOP2 overexpression mutants displayed the highest per cell total rAAV-DNA productivity (1.6 and 1.7-fold increase over

control, respectively) and per cell vector productivity (3 and 4-fold over control, respectively).

Conclusions: We identified proteins that positively impact rAAV2 vector production in a yeast high-throughput screening model of a subset of host cell factors. Additional analysis confirmed that overexpression of HEM4 and TOP2 proteins enhanced total and benzonase-resistant rAAV DNA yield. Further studies are needed to understand their mechanism of action and to assess their potential application in molecular strategies for rAAV production.

Keywords: Adeno-associated virus, vector production, *Saccharomyces*, yTHC screening

BACKGROUND.

Adeno-associated viruses (AAV) are single-stranded DNA viruses classified as *Dependoparvovirus*. AAV replication not only depends on viral and host cell factors but also on auxiliary viral proteins provided by a helper virus such as Herpesvirus or Adenovirus (Daya and Berns 2008). The AAV genome is ~4.7kb with three open reading frames, Rep, Cap, and AAP; flanked by inverted terminal repeats (ITR). The Rep gene encodes four non-structural proteins important for DNA replication and encapsidation (Rep40, 52, 68, and 78). The Cap gene encodes the three proteins that comprise the AAV icosahedral capsid (VP1, VP2, VP3). A third ORF overlapping the Cap gene encodes the assembly activating protein (AAP) that promotes VP shuttling into the nucleus and capsid assembly. The ITRs are cis-acting elements that contain recognition sequences important for DNA replication and packaging (Samulski and Muzyczka 2014). In-depth studies on AAV biology demonstrated that this virus could be used as a gene therapy vector due to its capacity to transduce specific genetic sequences into mammalian cells (Tratschin et al. 1984; Tratschin et al. 1985). Additional studies demonstrated long-term transgene expression, and clinical evaluations with recombinant AAV vectors showed very low inflammatory response and toxicity to patients, confirming rAAV's potential as a gene therapy vector (Carter 2005).

There are currently three production platforms available for industrial rAAV vector production. Adenovirus, Herpesvirus and Baculovirus complementation systems constitute robust platforms for vector production at pre-clinical and clinical scale (Ayuso et al. 2010; Aponte-Ubillus et al. 2017). Several process optimization studies achieved considerable improvement in vector particle formation by focusing on many bioengineering aspects such as cell growth, plasmid transfection, media optimization, and vector purification(Zolotukhin et al. 2002; Negrete and Kotin 2008; Thomas et al. 2009;

Potter et al. 2014; Robert et al. 2017). Vector volumetric yields range between 10^{10}-10^{11} vector genomes (vg) per mL of culture; however, the generation of sufficient rAAV quantities to supply drug commercialization remains a significant challenge. With the potential for very high clinical doses of 10^{14} vg/kg and limitation in cell culture scalability, several efforts are focused on improving per cell vector productivity and process scalability (Ayuso 2016; Clément and Grieger 2016). Molecular optimization of rAAV-producing cell lines has been studied extensively with significant attention to plasmid design, including the evaluation of different arrangements of AAV genes, promoters, and regulatory sequences (Xiao et al. 1998; Grimm et al. 1998; Allen et al. 2000). Because AAV is a *Dependoparvovirus*, substantial effort was oriented toward identifying the optimal helper genes necessary for vector production. Minimum subsets of Adenovirus and Herpesvirus helper genes were identified, and their functional analysis has been explored (Weindler and Heilbronn 1991; Matsushita et al. 1998; Geoffroy and Salvetti 2005). Interestingly, the role host cell factors play on wild type AAV and vector formation has been only partially investigated. Studies in mammalian cells led to the identification of AAV DNA replication complex components (Nash et al. 2007; Nash et al. 2008). In addition, protein-protein interaction analysis provided insight into the potential mechanisms DNA damage repair proteins exert on the AAV life cycle (Adachi and Nakai 2011). Current efforts focus on the virus-host cell interplay and how it modulates the AAV life cycle, with the objective of identifying molecular components that enhance vector particle formation.

S. cerevisiae is a very popular model organism due to its high growth rate, easy genetic manipulation, availability of multiple molecular tools, high-throughput capabilities, and documented transferability of findings into more complex organisms (Alves-Rodrigues et al. 2006; Nagy et al. 2014). Genetic and proteomic approaches have been utilized to study host-virus interactions in yeast models, primarily studying the fundamental biology of plant

and animal RNA viruses and DNA viruses (Zhao and Frazer 2002; Raghavan et al. 2004; Serviene et al. 2006; Jiang et al. 2006). Yeast has been recently used to explore basic aspects of rAAV particle formation. Alvaro Galli's group successfully demonstrated AAV single-stranded DNA replication and formation of AAV virus-like particles in *Saccharomyces cerevisiae* (Cervelli et al. 2011; Backovic et al. 2012; Galli et al. 2017). In addition, we have further developed a rAAV model system in yeast that recapitulated key aspects of AAV biology leading to formation of infectious rAAV particles (Barajas et al. 2017). In the current work we use *S. cerevisiae* as a model to study the effect of host cell factors on rAAV2 DNA replication and vector formation, ultimately identifying host proteins that improve rAAV vector production. By using the yeast Tet Hughes Collection (yTHC) screening library, we identified the HEM4 and TOP2 proteins as enhancers of cellular rAAV2-GFP productivity in *S. cerevisiae.*

METHODS.

Microorganisms and culture media

S. cerevisiae R1158 (URA3::CMV-tTA MATa his3-1 leu2-0 met15-0), a derivative of strain BY4741, was obtained from Dharmacon (GE Life Sciences) as part of the yTHC culture collection. *S. cerevisiae* YPH500 (MATα ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1) was obtained from Agilent Technologies. Glycerol stocks were maintained at -80ºC. YPD broth (yeast extract 1%, Peptone 2%, Dextrose 2%) was utilized for basic culture propagation. Yeast synthetic complete medium (SC) supplemented with 2% glucose and lacking the appropriate amino acids was used for plasmid transformation. To induce rAAV production, fresh SC medium with galactose (2-3%) was provided to the culture either by media exchange or feed addition after glucose depletion.

Plasmid design

A rAAV2-GFP producing plasmid system was originally described in Barajas *et al.* (Barajas et al. 2017), which comprised four pESC-2µ-based plasmids (DB205, DB155, DB029, DB040) that contain expression cassettes for six AAV proteins (Rep78, Rep52, AAP, VP1, VP2 and VP3) and an ITR-flanked GFP gene. A rAAV2-GFP 2-plasmid system was designed specifically for the screening of the yTHC library (Fig 1). All AAV coding regions of the previously mentioned 4-plasmid system were amplified and consolidated into JA001 plasmid by Gibson Assembly. Primers JA001-008 (see Additional File 1) were used to amplify four coding regions: His3-FKBP46-AAP-VP3 (from DB155), VP2-Rep78 (from DB029), 2µ-AmpR (from DB155), and VP1-E2A-Rep52 (from DB205). The JA002 plasmid was constructed by cloning the *Plu*TI/*Ngo*MIV fragment from pESC-LEU (Agilent Technologies) into the pAAV-GFP plasmid (Cell Biolabs), which was also digested with *Plu*TI and *Ngo*MIV. In addition, a pAAV-GFP-LYS2 plasmid was designed (DB320) as an alternative to JA002. The LYS2 gene from *S. cerevisiae* was amplified with primers DB633 and DB634, digested with *Pfo*I and ligated into DB040-pAAV-GFP-2µ-URA3 also digested with *Pfo*I.

To evaluate expression of individual proteins in the YPH500 rAAV-producing strain, pESC-URA overexpression plasmids were constructed. First, pESC-URA was modified to facilitate high throughput Gibson assembly. The GAL1/GAL10 promoter region was amplified with primers DB668 and DB669, digested with *BamHI* and *Not*I and ligated into *Bam*HI/*Not*I-digested pESC-URA, generating DB327-pESC(U)-2µ-GAL10. Second, the coding sequence of each protein of interest was amplified from *S. cerevisiae* genomic DNA using specific primers JA009-JA018 (see Additional File 1), and the amplicon was

later inserted by Gibson assembly into *Sma*I-linearized DB327, downstream of GAL10 promoter.

yTHC library screening

The yTHC screening library was acquired from Dharmacon (GE Life Sciences). This commercial library contains more than 800 yeast mutants with individual integration of a kanR-Tet07-TATA cassette into the genome, replacing the natural promoter of specific essential genes. Tet promoter strength is higher than many yeast endogenous promoters, which in many cases leads to gene overexpression in the absence of Doxycicline (Dox). Addition of Dox 10µg/mL, on the other hand, switches off the promoter activity (Mnaimneh et al. 2004; Serviene et al. 2006).

yTHC strains were transformed with plasmids JA001 and JA002 using a lithium acetate/polyethylene glycol (LiAc/PEG) method. Transformed clones were selected by complementation of auxotrophic markers using SC medium lacking histidine and leucine (HL^-). Two transformed clones of each strain were evaluated for rAAV production. Briefly, each clone was initially propagated in SC HL^- + 2% glucose at 30°C and 825 rpm of agitation, for 36 hours using 2-mL 96 deep well plates (Axygen) and an orbital plate shaker (VWR). A split ratio of approximately 1:10 was used to inoculate a subsequent plate containing 750µL/well of fresh SC HL^- media broth plus 2% glucose, reaching an initial optical density of ~0.5 OD600nm. The cultures were grown for 24 hours at the aforementioned conditions. Immediately afterwards, media exchange with galactose-enriched (2%) SC medium was performed to induce expression of AAV proteins. Doxycycline (10µg/mL) was added, when appropriate, immediately following media

exchange. The culture was incubated for 96 hours after induction, maintaining the same operating parameters.

Total rAAV-GFP DNA, 18s ribosomal DNA and benzonase-resistant rAAV-GFP DNA (also referred as vector yield) were measured by digital droplet PCR (ddPCR). Total rAAV-GFP DNA values were normalized to 18S ribosomal DNA values and shown as a measure of replicated rAAV DNA (rAAV-GFP/18s rDNA ratio). Details are explained in a subsequent section. The average values for each result were compared relative to the transformed parental strain R1158, and potential lead strains were selected if vector yield or rAAV-GFP/18s rDNA ratio surpassed the control performance values by 2-fold and 0.5-fold, respectively. Positive strains were confirmed in a subsequent screening round. Three pools of clones from each strain were fermented in 24 deep well plates at Dox +/- conditions, using the same operating parameters described above with the exception that agitation was changed to 550 rpm.

Lead validation

Selected protein candidates identified in the yTHC screening were overexpressed in *S. cerevisiae* YPH500. This strain was initially transformed with plasmids DB155, DB145, DB320 and DB29 to produce rAAV. The strain was subsequently transformed with DB327-pESC(U)-2µ-GAL10 plasmids carrying the coding sequence for the protein of interest. Four clones of each strain were evaluated in 24 deep-well plates, using the method described above. Unstable clones that lost one or more plasmids during fermentation were removed from the analysis. Cell lysates were taken to perform ddPCR rAAV DNA analysis. Lysate was further purified over an AVB affinity resin (GE Healthcare) and analyzed for capsid titer analysis and capsid-to-vector genome (cp:vg) ratio determination.

To further confirm the positive effect of a selected overexpression mutant, a Tet-integrated, YPH500-TOP2 strain was generated. A KanMX-TETp cassette was integrated upstream of the TOP2 coding sequence by homologous recombination. The G418 resistance cassette KanMX was amplified by PCR from plasmid pUG6-G418 (Güldener et al. 1996) using primers DB560 and DB561. The PCR product was digested with *Sal*I and inserted into plasmid pCM190-TETp (Garí et al. 1997) linearized with *Xho*I. The resulting plasmid was used as a template for PCR with primers DB574 and DB575 to generate the KanMX-TETp cassette flanked by 40 nucleotides with homology to the target site upstream of the TOP2 coding sequence. *S. cerevisiae* YPH500 was transformed with the KanMX-TETp PCR product, plated on YPD for 24 hours and streaked onto YPD supplemented with 100 mg/L G418. Integration of the KanMX-TETp cassette in the target site was confirmed with primers DB471 and DB562. Selected clones were transformed with plasmids DB205, DB155, DB029 and DB040 to produce rAAV.

Bioreactor runs

The top clones from the best overexpression candidates were evaluated in bioreactor runs against the control strain YPH500 clone. Fermentations were performed using a Dasbox 250mL mini-bioreactor system (Eppendorf) at 30ºC, pH 4.5, 30% dissolved oxygen (DO) by air saturation, and 0.5 vvm aeration. The aeration loop involved integrated increase in agitation (300-1000rpm), total flow rate (6-12sL/h) and pure oxygen flow rate increase in response to low DO values. Fermentations were operated as fed-batch mode. Briefly, 200mL of glucose-based, SC medium pH 5.5 supplemented with phosphate buffer to 0.1M was inoculated with a fresh culture at initial density of 0.15 OD600nm. A 10X bolus feed of galactose-based SC medium was added to the bioreactor when glucose concentration fell below 1g/L. Residual glucose and galactose concentrations were monitored

throughout the subsequent 96 hours via the Cedex Bio HT metabolite analyzer (Roche). Optical cell density and dry cell weight was determined to establish growth and productivity rates. At termination, cells were harvested and lysates were used for quantification of total and benzonase-resistant rAAV-GFP DNA by ddPCR. AVB affinity purified capsids were quantified and the cp:vg ratio was calculated by dividing the capsid ELISA titer (assay description below) by the benzonase resistant rAAV-GFP ddPCR vg titer. Productivity ratios were also determined at the end of the run. Per cell vector productivity was determined by dividing day 5 rAAV-GFP vector genome (vg) titer by final biomass. Per cell AAV DNA replication was determined by dividing total rAAV DNA on day 5 by final biomass.

rAAV purification and analytical testing

Yeast cultures were collected by centrifugation at 4000rpm x 3min. Cells were resuspended in 10 mM Tris-HCl, 1 mM EDTA buffer supplemented with 1mg/mL Zymolase and 0.1%v/v beta-mercaptoethanol; and incubated for 1 hour at 37°C to digest the cell wall. Approximately 10 µL of each cell suspension were aliquoted into 300 µL of 1% SDS, incubated for 10 minutes at 60°C and sonicated for 3 minutes in a sonication bath prior to sample dilution for total rAAV DNA and 18S rDNA analysis. The remaining cell lysate suspension was treated with an equal volume of a nuclease buffer containing 15U/mL benzonase, 250mM NaCl, 0.5% sarkosyl, 0.5% Triton X-100, 0.05% Pluronic-F68 and 5mM $MgCl_2$. The suspension was incubated for 2 hours at 37ºC to digest unprotected rAAV DNA. Cell debris was removed by centrifugation at 100g x 2min. The clarified lysate was used for vector yield quantification (benzonase-resistant rAAV-GFP DNA) and for total capsid purification.

High-throughput ddPCR was used to quantify total rAAV-GFP DNA and 18s rDNA from cell lysate to determine average rAAV-GFP DNA copies per cell (rAAV-GFP DNA/18s rDNA ratio). To determine the amount of benzonase-resistant rAAV-GFP DNA, benzonase-treated lysate was purified using 50µL AVB sepharose resin (GE) on 96 deep-well plates as previously described (Barajas et al. 2017), and the rAAV-GFP DNA copy number from AVB eluates was analyzed by ddPCR.

Digital droplet PCR was performed as described previously (Barajas et al. 2017). Yeast treated material was diluted 10-to-100-fold to target the ddPCR dynamic range. Five µL of diluted material were mixed with 20 µL of Taqman-based master mix (BioRad) including GFP and 18s rDNA primers and probes (DB307/309 and DB438/440 respectively), using FAM and HEX reporting dyes DB308 and DB439. Droplets were generated by an automated droplet generator (Biorad), and amplified material was analyzed in a QX200 droplet reader (Biorad) using Quantasoft software (Biorad).

For mRNA analysis, 400uL of pellet samples were mechanically disrupted using a Fastprep24 homogenizer, using 3x 30-second cycles with intermediate rest times. Supernatant was aliquoted and total RNA was extracted using RNAeasy kit (Qiagen) following the manufacturer's guidelines. RNA concentration was measured via Nanodrop. 100ng of RNA were used for a reverse-transcription (RT) reaction using iScript cDNA synthesis kit (BioRad). The subsequent cDNA was diluted 1:1 and 1:10 and quantified by ddPCR using specific primers and probes for host cell and Rep sequences.

Capsid ELISA was performed to quantify AAV2 capsid yield. Yeast pellet (200uL) was zymolase/benzonase-treated as described above, and rAAV capsids were isolated by AVB resin purification. Three-fold dilutions of the purified material were analyzed with Progen AAV2 titration kit, following the manufacturer's instructions. Southern and western

blot were performed to characterize rAAV-GFP DNA and VP protein expression in yeast. As previously described (Barajas et al. 2017).

RESULTS.

yTHC screening identifies genes that impact rAAV production

A high-throughput screening strategy was designed to evaluate rAAV2-GFP replication in yeast strains where the expression of essential genes is modulated. Yeast growing conditions were adapted to generate sufficient yeast biomass and detectable levels of total and benzonase-resistant rAAV2 DNA. The original AAV expression cassettes described in Barajas (Barajas et al. 2017) were arranged into one plasmid to meet yTHC library auxotrophic genetic requirements and generate rAAV DNA replication and VP expression (Fig 1).

Analysis of AVB-purified and non-purified cell lysate material from more than 800 plasmid-transformed, yTHC strains led to the identification of 22 strains with increased rAAV DNA and vector yield. Analysis of rAAV-GFP/18s rDNA ratio showed moderate improvement (1-2 fold change relative to control) in 18 strains, whereas 4 strains (*RRN3, GPN3, RRP14, PRE4*) had strong improvements as high as 6-fold in -Dox conditions (Fig 2A, 2B). ddPCR analysis of benzonase-resistant AAV-GFP DNA showed moderate enhancement of vector yield (1-4 fold change relative to control) in 18 strains, and strong enhancement in 4 strains (*SDO1,TOP2, HEM4, PRE4*) reaching values as high as 16-fold relative to control strain.

To prevent a negative impact in biomass formation and rAAV yield as a result of downregulation of essential genes, we performed the molecular screening at growing conditions that resemble current yeast bioprocesses. When strains were subjected to a 24 hour growth phase prior to Dox addition time, rAAV yield trends were similar in both

Dox presence and absence conditions (data not shown). These yield improvements correlated with mRNA transcriptional upregulation of the gene of interest (Fig. 3). Rep mRNA levels were analyzed and showed minimal differences among conditions. These results prompted us to evaluate the dynamics of transcriptional downregulation triggered by Dox addition under these specific growing conditions. We took the TOP2 strain as a case study. Under this evaluation, cells were cultured similarly to the 96-well plate screening protocol; however, they were inoculated at an initial density of 0.05 OD600nm. Such low inoculation density was required to see the effect of early Dox addition times. mRNA levels were analyzed at galactose induction time. Under Dox-free conditions, *TOP2* mRNA levels in the yTHC-TOP2 strain were 60-fold higher than in the R1158 control strain (Fig 4). Dox addition at 0 hours and 16 hours before galactose induction reduced the mRNA fold change to 3-fold and 0.8-fold relative to the control strain, respectively. Direct correlation between *TOP2* mRNA levels and benzonase-resistant AAV-GFP DNA suggested this protein is exerting an enhancing effect on rAAV production. Overall, the previous data support three concepts: 1) Significant gene knockdown/downregulation was not accomplished during the screening; 2) yeast growing conditions used in this study may be suboptimal for Dox-mediated gene downregulation assays; 3) improvement of rAAV yield might have occurred as a result of TOP2 overexpression.

Overexpression mutants confirm rAAV-promoting effect of selected host proteins.

The aforementioned results suggested that rAAV yield improvement seen in selected yeast strains might be linked to host protein overexpression. To confirm that hypothesis, we overexpressed the best 5 candidate proteins from the screening library (PRE4, HEM4, TOP2, SDO1, and GPN3) using a different haploid yeast strain carrying 4 plasmids to produce rAAV as described in Barajas *et al* (Barajas et al. 2017). Analysis of triplicate

colonies of each YPH500 strain in 24 deep well plates corroborated the results seen during yTHC screening. HEM4, GPN3 and SDO1 overexpression strains showed 40% improvement in AAV DNA replication, whereas TOP2 overexpression strain exhibited 80% improvement. Analysis of benzonase-resistant AAV DNA in yeast lysate had low to moderate increases in the SDO1, GPN3, HEM4 strains, ranging from 10-80%. The TOP2 overexpression strain presented the best performance, improving vector yield 650% relative to control (Fig 5A, 5B).

To further confirm the rAAV-enhancing activity of the most promising host protein candidate, an additional evaluation of TOP2 overexpression was performed by integrating a Tet-regulatable promoter upstream the TOP2 coding sequence, in a YPH500 yeast strain. Different clones resulting from Tet-TOP2 cassette integration and rAAV plasmid transformation were evaluated. Productivity of these Tet-TOP2 strains aligned with our plasmid overexpression studies. The best clone generated rAAV2 vector yield of $9.5x10^8$ vg/mL, 8.5-fold higher than the 4-plasmid, YPH500 control strain (Fig. 5C). This data confirmed overexpression results seen during the preliminary screening.

HEM4 and TOP2 mutants improve rAAV vector yield under bioreactor-controlled conditions.

To corroborate consistent rAAV productivity of the most promising yeast strains, the performance of the best HEM4, TOP2 and GPN3 overexpression clones under controlled stirred-tank fermentation conditions was assessed. Dry cell weight and OD600nm values showed 25% lower biomass formation in TOP2 and GPN3 strains, compared to the YPH500 control strain. However, growth trends were comparable for all strains (Fig 6A, 6B). As expected, the highest growth rate was achieved during the glucose biomass

expansion phase and then slowed during the galactose induction phase. Glucose was depleted to less than 1g/L approximately 24 hours after inoculation. At this time, galactose feed was added to induce rAAV protein expression. Thirty grams per liter of galactose were sufficient to support continuous growth for most strains during the last 4 days of culture. The HEM4 strain required an extra 5g/L of galactose to prevent carbon source depletion.

Cell lysate rAAV DNA analysis showed similar trends of improvement for HEM4 and TOP2 overexpression strains, reaching vector yields (benzonase-resistant rAAV-GFP titer) ranging 7-8 x 10^8 vg/mL and total rAAV DNA yields of 7-9 x 10^{10} vg/mL (Fig 6C, 6D). Contrary to TOP2 and HEM4, the GPN3 strain had rAAV DNA levels slightly lower than control. The HEM4 and TOP2 strains had improved per cell productivity of 3-fold along with a 60-70% increase in total rAAV DNA. Furthermore, the ratio of capsids to vector genomes dropped as much as 50% for the TOP2 mutant, demonstrating improved DNA encapsidation (Fig 6E). Overall, these results suggest that yeast rAAV-GFP production is enhanced by overexpression of host proteins HEM4 and TOP2.

DISCUSSION.

In this study, a *Saccharomyces* rAAV production model was used in combination with the yTHC commercial library to screen for host factors that enhance rAAV DNA and rAAV-GFP full particle production. The 22 protein candidates identified by this screen participate in biological processes including RNA processing, DNA replication, ribosome biogenesis and protein processing. Subsequent confirmatory screens identified TOP2 and HEM4 overexpression as drivers of improved rAAV production in a yeast model.

Several groups have studied the host protein network responsible for AAV formation, leading to the identification and partial functional determination of several proteins. Nash and collaborators identified host proteins MCM, PCNA, POLδ, RFA, and RFC as part of the AAV DNA replication fork(Nash et al. 2008) and were able to reproduce AAV DNA replication *in vitro*. Results from this study pointed out that cellular fractions were still needed to promote viral replication, suggesting that additional proteins were required. Proteomic approaches identified further host components that interact with Rep proteins and that are likely part of AAV replication compartments (Nash et al. 2009; Nicolas et al. 2010). These studies highlighted the multi-functionality of Rep proteins, as their interaction network included host proteins that participate in a plethora of biological processes. Specific studies targeting individual host proteins confirmed the capacity of mammalian proteins such as Topors, PC4, and ANP32B, to increase viral replication in different *in vivo* and *in vitro* models (Muramatsu et al. 1998; Weger et al. 2002; Pegoraro et al. 2006).

Gene candidates identified by the yTHC library (*HEM4, TOP2, SDO1*, *PRE4* and *GPN3*) were further screened by overexpression in YPH500 yeast. Notably, YPH500 PRE4 overexpression did not improve rAAV vector concentrations compared to control strain, suggesting that the potential mechanisms promoting AAV production were strain-specific. Similarly, the rAAV-enhancing capacity of the GPN3 overexpression mutant was inconsistent across different culture conditions (controlled bioreactor vs. deep-well plate), which indicates that the positive impact of this protein could presumably be linked to physiological responses to culture conditions such as low pH, low oxygen levels, or high byproduct formation. It should also be considered that lower vector productivity was evidenced in the 2-plasmid rAAV producing system, reason that could explain the less pronounced yield improvement when the selected host proteins where overexpressed in the 4-plasmid rAAV system.

However, the HEM4 and TOP2 mutants showed consistent vector productivity across all experimental procedures. Per cell vector productivity (rAAV-GFP vector titer/ g DCW) in both mutants was improved by 3-4 fold relative to control, and capsid to vg ratio analysis suggests that this enhancement is in part due to increased encapsidation. HEM4 encodes uroporphyrogen III synthase, an enzyme necessary for the biosynthesis of tetrapyrrole-based compounds such as heme (Saccharomyces Genome Database: https://www.yeastgenome.org/). Heme is an important cofactor for several biological processes, especially respiration. High heme levels correlate with high mitochondrial respiration and activation of oxidative stress response, repression of anaerobic metabolism (Martínez et al. 2016a), increased heterologous protein production, and decreased expression of stress-related transcription factor Hap1p(Martínez et al. 2016b). Any or all of these mechanisms could plausibly have a positive effect on AAV production.

TOP2 encodes the yeast topoisomerase II. This multifunctional protein relieves topological tension in DNA throughout different biological processes such as DNA replication, transcription and chromosome maintenance (Nitiss 2009). TOP2 contributes to DNA genome stability and ameliorates potential DNA replication and transcription stress effects (Bermejo et al. 2009; Pannunzio and Lieber 2016). Topoisomerase II activity has a positive impact in replication of other DNA viruses such as SV40 and Adenovirus, presumably by maintaining stability of double-stranded viral DNA and preventing concatemerization of synthesized genetic material (Wong and Hsu 1990; Ishimi et al. 1992). We speculate that overexpression of HEM4 and TOP2 proteins in yeast might cope with detrimental effects potentially resulting from rAAV formation, such as genetic instability caused by AAV Rep78-DNA replication activities, or proteotoxic and oxidative stress due to protein overexpression. A functional analysis of these host elements will be key to elucidate the specific biological implications of HEM4 and TOP2 on rAAV production. Mammalian

Top2α/β and insect Top2 have very conserved functions relative to the yeast homologue. This is also true among mammalian UROS, insect CG1885/9589 and yeast HEM4 proteins. Further studies to evaluate the transferability of the information generated in this yeast screening model would assess whether the animal homologues of TOP2 and HEM4 proteins exert a similar effect in traditional rAAV production systems such as HEK293 and Sf9 cells.

CONCLUSIONS

The rAAV2-GFP producing yeast system is a promising model for the evaluation of the impact of host cell factors on rAAV production. Results shown in Barajas *et al* (Barajas et al. 2017) as well as in this manuscript indicate that the yeast model is able to recapitulate key aspects of AAV formation such as DNA replication and full particle formation. Our results complement previous studies done in yeast and support the notion of the utility of this model. Due to its simplicity, adaptability to high-throughput formats, and fast turnaround time; this system could become a valuable tool for other scientists interested in study critical aspects of rAAV generation in eukaryotic cells.

Using high-throughput screening we identified host cell factors that positively impact rAAV2 vector production in a *S. cerevisiae* model. Additional analysis confirmed that overexpression of HEM4 and TOP2 proteins enhanced total and benzonase-resistant rAAV DNA yields. These results contribute to the understanding of rAAV-producing systems and to the biotechnological development of microbial platforms for rAAV vector production.

DECLARATIONS

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material: The datasets supporting the conclusions of this article are included within the article.

Funding: This study was funded by Biomarin Pharmaceutical Inc.

Authors' contributions: All authors participated in the conceptualization of the study. JJAU, DB, JP and CB contributed to the experimental design and data analysis. JJAU and DB performed the investigation. JJAU wrote the manuscript and all authors contributed to its edition.

Competing interests: DB, JP and DG are employees and stockholders of Biomarin Pharmaceutical Inc.

Acknowledgments: The authors would like to thank Javier Femenia, Yvette Tang and Tomas Cinek at BioMarin Pharmaceutical Inc for their valuable suggestions and support throughout the execution of the study.

REFERENCES.

1. Adachi K, Nakai H (2011) The Role of DNA Repair Pathways in Adeno-Associated Virus Infection and Viral Genome Replication / Recombination / Integration. In: Vengrova S (ed) DNA Repair and Human Health. InTech

2. Allen JM, Halbert CL, Miller AD (2000) Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6. Mol Ther J Am Soc Gene Ther 1:88–95 . doi: 10.1006/mthe.1999.0010

3. Alves-Rodrigues I, Galão RP, Meyerhans A, Díez J (2006) Saccharomyces cerevisiae: a useful model host to study fundamental biology of viral replication. Virus Res 120:49–56 . doi: 10.1016/j.virusres.2005.11.018

4. Aponte-Ubillus JJ, Barajas D, Peltier J, Bardliving C, Shamlou P, Gold D (2017) Molecular design for recombinant adeno-associated virus (rAAV) vector production. Appl Microbiol Biotechnol. doi: 10.1007/s00253-017-8670-1

5. Ayuso E (2016) Manufacturing of recombinant adeno-associated viral vectors: new technologies are welcome. Mol Ther Methods Clin Dev 3:15049 . doi: 10.1038/mtm.2015.49

6. Ayuso E, Mingozzi F, Bosch F (2010) Production, purification and characterization of adeno-associated vectors. Curr Gene Ther 10:423–436

7. Backovic A, Cervelli T, Salvetti A, Zentilin L, Giacca M, Galli A (2012) Capsid protein expression and adeno-associated virus like particles assembly in Saccharomyces cerevisiae. Microb Cell Factories 11:124 . doi: 10.1186/1475-2859-11-124

8. Barajas D, Aponte-Ubillus JJ, Akeefe H, Cinek T, Peltier J, Gold D (2017) Generation of infectious recombinant Adeno-associated virus in Saccharomyces cerevisiae. PloS One 12:e0173010 . doi: 10.1371/journal.pone.0173010

9. Bermejo R, Capra T, Gonzalez-Huici V, Fachinetti D, Cocito A, Natoli G, Katou Y, Mori H, Kurokawa K, Shirahige K, Foiani M (2009) Genome-organizing factors Top2 and Hmo1 prevent chromosome fragility at sites of S phase transcription. Cell 138:870–884 . doi: 10.1016/j.cell.2009.06.022

10. Carter BJ (2005) Adeno-Associated Virus Vectors in Clinical Trials. Hum Gene Ther 16:541–550 . doi: 10.1089/hum.2005.16.541

11. Cervelli T, Backovic A, Galli A (2011) Formation of AAV Single Stranded DNA Genome from a Circular Plasmid in Saccharomyces cerevisiae. PLOS ONE 6:e23474 . doi: 10.1371/journal.pone.0023474

12. Clément N, Grieger JC (2016) Manufacturing of recombinant adeno-associated viral vectors for clinical trials. Mol Ther Methods Clin Dev 3:16002 . doi: 10.1038/mtm.2016.2

13. Daya S, Berns KI (2008) Gene Therapy Using Adeno-Associated Virus Vectors. Clin Microbiol Rev 21:583–593 . doi: 10.1128/CMR.00008-08

14. Galli A, Della Latta V, Bologna C, Pucciarelli D, Cipriani F, Backovic A, Cervelli T (2017) Strategies to optimize capsid protein expression and single-stranded DNA formation of adeno-associated virus in Saccharomyces cerevisiae. J Appl Microbiol 123:414–428 . doi: 10.1111/jam.13511

15. Garí E, Piedrafita L, Aldea M, Herrero E (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. Yeast Chichester Engl 13:837–848 . doi: 10.1002/(SICI)1097-0061(199707)13:9<837::AID-YEA145>3.0.CO;2-T

16. Geoffroy M-C, Salvetti A (2005) Helper functions required for wild type and recombinant adeno-associated virus growth. Curr Gene Ther 5:265–271

17. Grimm D, Kern A, Rittner K, Kleinschmidt JA (1998) Novel Tools for Production and Purification of Recombinant Adenoassociated Virus Vectors. Hum Gene Ther 9:2745–2760 . doi: 10.1089/hum.1998.9.18-2745

18. Güldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH (1996) A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Res 24:2519–2524

19. Ishimi Y, Ishida R, Andoh T (1992) Effect of ICRF-193, a novel DNA topoisomerase II inhibitor, on simian virus 40 DNA and chromosome replication in vitro. Mol Cell Biol 12:4007–4014

20. Jiang Y, Serviene E, Gal J, Panavas T, Nagy PD (2006) Identification of Essential Host Factors Affecting Tombusvirus RNA Replication Based on the Yeast Tet Promoters Hughes Collection. J Virol 80:7394–7404 . doi: 10.1128/JVI.02686-05

21. Martínez JL, Meza E, Petranovic D, Nielsen J (2016a) The impact of respiration and oxidative stress response on recombinant α-amylase production by Saccharomyces cerevisiae. Metab Eng Commun 3:205–210 . doi: 10.1016/j.meteno.2016.06.003

22. Martínez JL, Petranovic D, Nielsen J (2016b) Heme metabolism in stress regulation and protein production: From Cinderella to a key player. Bioengineered 7:112–115 . doi: 10.1080/21655979.2015.1126016

23. Matsushita T, Elliger S, Elliger C, Podsakoff G, Villarreal L, Kurtzman GJ, Iwaki Y, Colosi P (1998) Adeno-associated virus vectors can be efficiently produced without helper virus. Gene Ther 5:938–945 . doi: 10.1038/sj.gt.3300680

24. Mnaimneh S, Davierwala AP, Haynes J, Moffat J, Peng W-T, Zhang W, Yang X, Pootoolal J, Chua G, Lopez A, Trochesset M, Morse D, Krogan NJ, Hiley SL, Li Z, Morris Q, Grigull J, Mitsakakis N, Roberts CJ, Greenblatt JF, Boone C, Kaiser CA, Andrews BJ, Hughes TR (2004) Exploration of essential gene functions via titratable promoter alleles. Cell 118:31–44 . doi: 10.1016/j.cell.2004.06.013

25. Muramatsu S, Handa A, Kajigaya S, Brown KE (1998) Transcription-positive cofactor 4 enhances rescue of adeno-associated virus genome from an infectious clone. J Gen Virol 79 (Pt 9):2157–2161 . doi: 10.1099/0022-1317-79-9-2157

26. Nagy PD, Pogany J, Lin J-Y (2014) How yeast can be used as a genetic platform to explore virus-host interactions: from “omics” to functional studies. Trends Microbiol 22:309–316 . doi: 10.1016/j.tim.2014.02.003

27. Nash K, Chen W, McDonald WF, Zhou X, Muzyczka N (2007) Purification of host cell enzymes involved in adeno-associated virus DNA replication. J Virol 81:5777–5787 . doi: 10.1128/JVI.02651-06

28. Nash K, Chen W, Muzyczka N (2008) Complete In Vitro Reconstitution of Adeno-Associated Virus DNA Replication Requires the Minichromosome Maintenance Complex Proteins. J Virol 82:1458–1464 . doi: 10.1128/JVI.01968-07

29. Nash K, Chen W, Salganik M, Muzyczka N (2009) Identification of cellular proteins that interact with the adeno-associated virus rep protein. J Virol 83:454–469 . doi: 10.1128/JVI.01939-08

30. Negrete A, Kotin RM (2008) Strategies for manufacturing recombinant adeno-associated virus vectors for gene therapy applications exploiting baculovirus technology. Brief Funct Genomic Proteomic 7:303–311 . doi: 10.1093/bfgp/eln034

31. Nicolas A, Alazard-Dany N, Biollay C, Arata L, Jolinon N, Kuhn L, Ferro M, Weller SK, Epstein AL, Salvetti A, Greco A (2010) Identification of Rep-Associated Factors in Herpes Simplex Virus Type 1-Induced Adeno-Associated Virus Type 2 Replication Compartments. J Virol 84:8871–8887 . doi: 10.1128/JVI.00725-10

32. Nitiss JL (2009) DNA topoisomerase II and its growing repertoire of biological functions. Nat Rev Cancer 9:327–337 . doi: 10.1038/nrc2608

33. Pannunzio NR, Lieber MR (2016) Dissecting the Roles of Divergent and Convergent Transcription in Chromosome Instability. Cell Rep 14:1025–1031 . doi: 10.1016/j.celrep.2015.12.098

34. Pegoraro G, Marcello A, Myers MP, Giacca M (2006) Regulation of adeno-associated virus DNA replication by the cellular TAF-I/set complex. J Virol 80:6855–6864 . doi: 10.1128/JVI.00383-06

35. Potter M, Lins B, Mietzsch M, Heilbronn R, Van Vliet K, Chipman P, Agbandje-McKenna M, Cleaver BD, Clément N, Byrne BJ, Zolotukhin S (2014) A simplified purification protocol for recombinant adeno-associated virus vectors. Mol Ther Methods Clin Dev 1:14034 . doi: 10.1038/mtm.2014.34

36. Raghavan V, Malik PS, Choudhury NR, Mukherjee SK (2004) The DNA-A Component of a Plant Geminivirus (Indian Mung Bean Yellow Mosaic Virus) Replicates in Budding Yeast Cells. J Virol 78:2405–2413 . doi: 10.1128/JVI.78.5.2405-2413.2004

37. Robert M-A, Chahal PS, Audy A, Kamen A, Gilbert R, Gaillet B (2017) Manufacturing of recombinant adeno-associated viruses using mammalian expression platforms. Biotechnol J 12:1600193 . doi: 10.1002/biot.201600193

38. Samulski RJ, Muzyczka N (2014) AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. Annu Rev Virol 1:427–451 . doi: 10.1146/annurev-virology-031413-085355

39. Serviene E, Jiang Y, Cheng C-P, Baker J, Nagy PD (2006) Screening of the yeast yTHC collection identifies essential host factors affecting tombusvirus RNA recombination. J Virol 80:1231–1241 . doi: 10.1128/JVI.80.3.1231-1241.2006

40. Thomas DL, Wang L, Niamke J, Liu J, Kang W, Scotti MM, Ye G-J, Veres G, Knop DR (2009) Scalable Recombinant Adeno-Associated Virus Production Using Recombinant Herpes Simplex Virus Type 1 Coinfection of Suspension-Adapted Mammalian Cells. Hum Gene Ther 20:861–870 . doi: 10.1089/hum.2009.004

41. Tratschin JD, Miller IL, Smith MG, Carter BJ (1985) Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. Mol Cell Biol 5:3251–3260

42. Tratschin JD, West MH, Sandbank T, Carter BJ (1984) A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. Mol Cell Biol 4:2072–2081

43. Weger S, Hammer E, Heilbronn R (2002) Topors, a p53 and topoisomerase I binding protein, interacts with the adeno-associated virus (AAV-2) Rep78/68 proteins and enhances AAV-2 gene expression. J Gen Virol 83:511–516 . doi: 10.1099/0022-1317-83-3-511

44. Weindler FW, Heilbronn R (1991) A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. J Virol 65:2476–2483

45. Wong ML, Hsu MT (1990) Involvement of topoisomerases in replication, transcription, and packaging of the linear adenovirus genome. J Virol 64:691–699

46. Xiao X, Li J, Samulski RJ (1998) Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. J Virol 72:2224–2232

47. Zhao K-N, Frazer IH (2002) Saccharomyces cerevisiae Is Permissive for Replication of Bovine Papillomavirus Type 1. J Virol 76:12265–12273 . doi: 10.1128/JVI.76.23.12265-12273.2002

48. Zolotukhin S, Potter M, Zolotukhin I, Sakai Y, Loiler S, Fraites TJ, Chiodo VA, Phillipsberg T, Muzyczka N, Hauswirth WW, Flotte TR, Byrne BJ, Snyder RO (2002) Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors. Methods San Diego Calif 28:158–167

49. HEM4 | SGD. https://www.yeastgenome.org/locus/HEM4. Accessed 8 Sep 2017

FIGURES

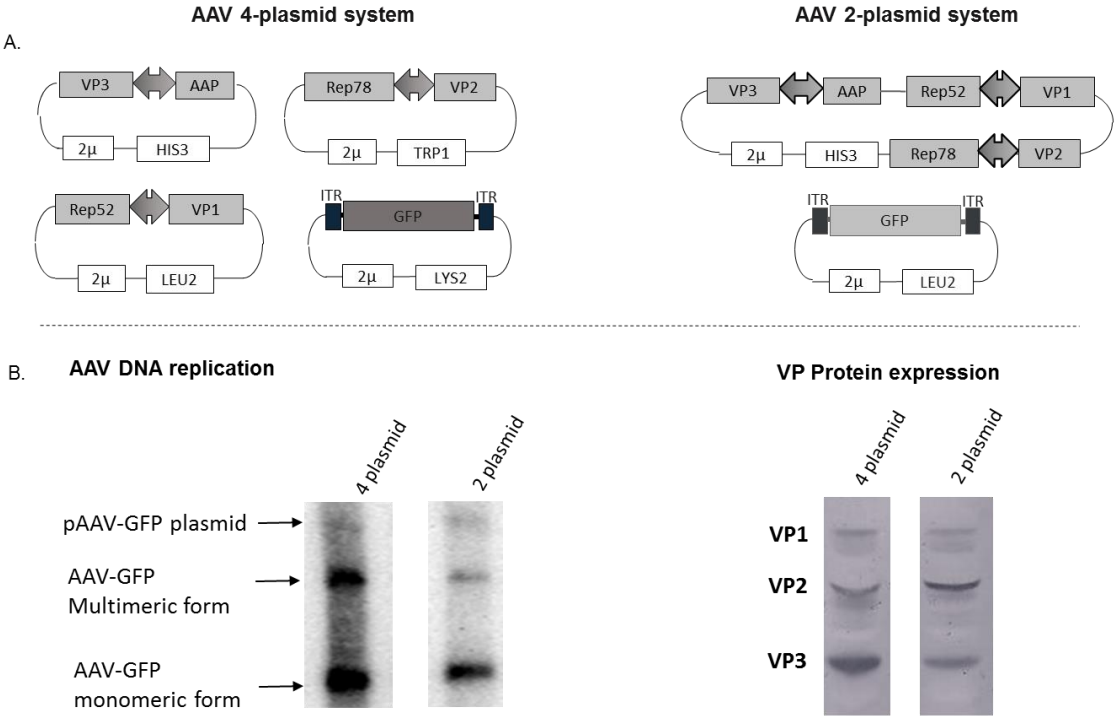
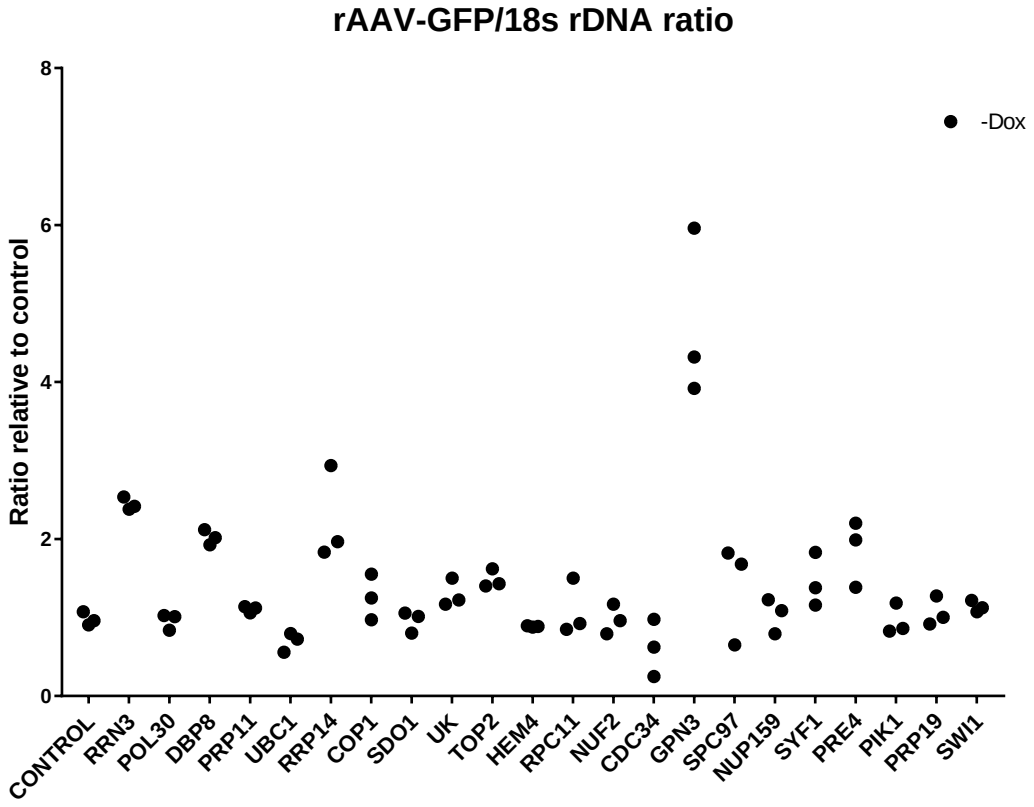


Fig 1. (A) rAAV-producing plasmid systems used in this study. (B) Evidence of AAV DNA replication and VP monomer expression in both systems.



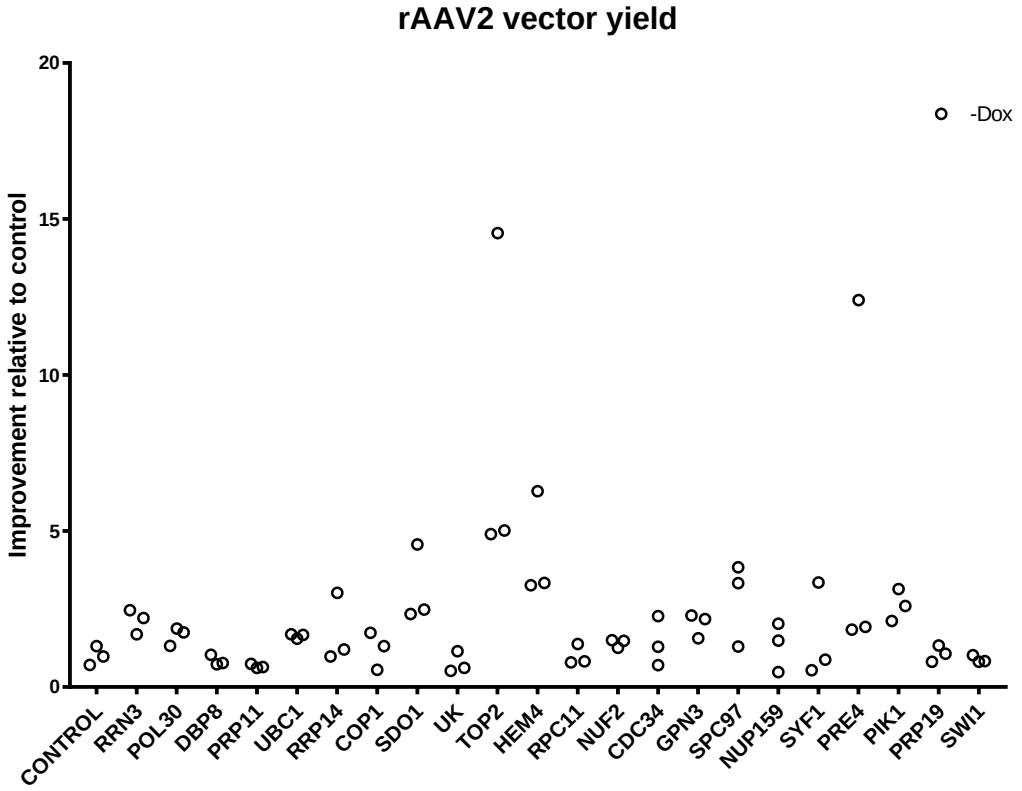


Fig 2. rAAV ddPCR analysis of 22 selected yTHC strains transformed with AAV 2-plasmid system. AAV-GFP/18s rDNA ratio analysis (top panel) and benzonase-resistant AAV DNA analysis (bottom panel) was performed using yeast lysate from 3 biological replicates of each mutant strain. Values are displayed as relative to control strain R1158 transformed with AAV 2-plasmid system.

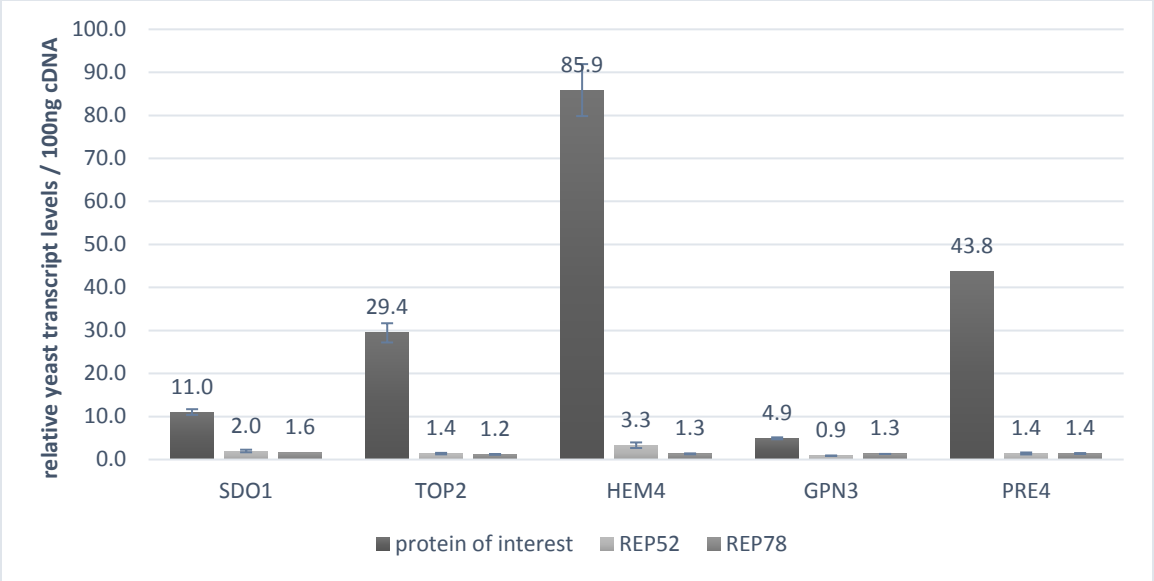


Fig 3. Host mRNA transcriptional levels of selected minus Dox yTHC strains relative to control strain. 100 ng samples were taken on day 3 post induction for analysis. Results are displayed as relative to control strain R1158.

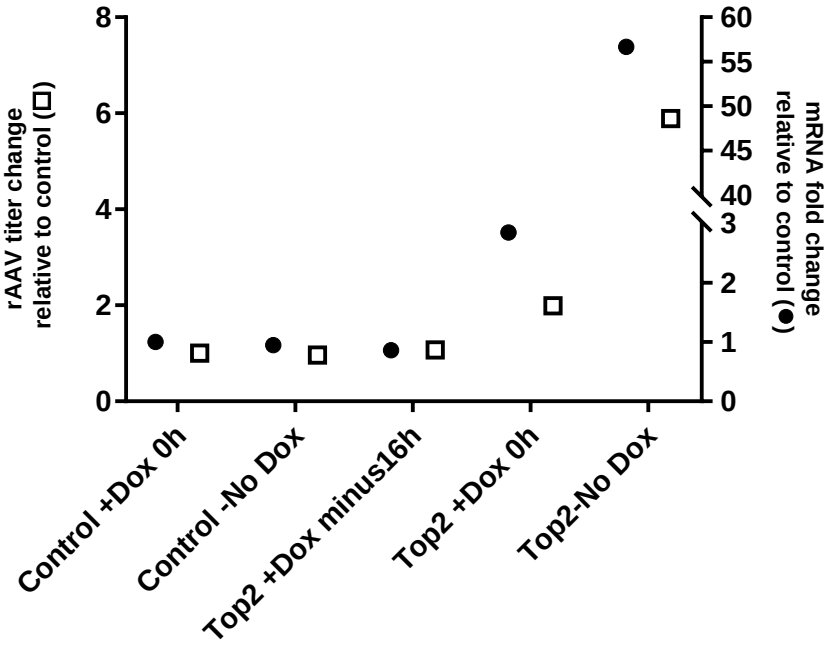


Fig 4. Effect of Doxycycline addition times on TOP2 mRNA transcriptional levels and vector yield. Results for yTHC TOP2 strain are shown relative to control R1158 strain. Dox was added as shown in hours post-galactose induction. 300 µL pellet samples were taken 2 hours after galactose induction for mRNA analysis. 1mL pellet samples were taken on day 4 post galactose induction to perform benzonase-resistant AAV DNA quantification.

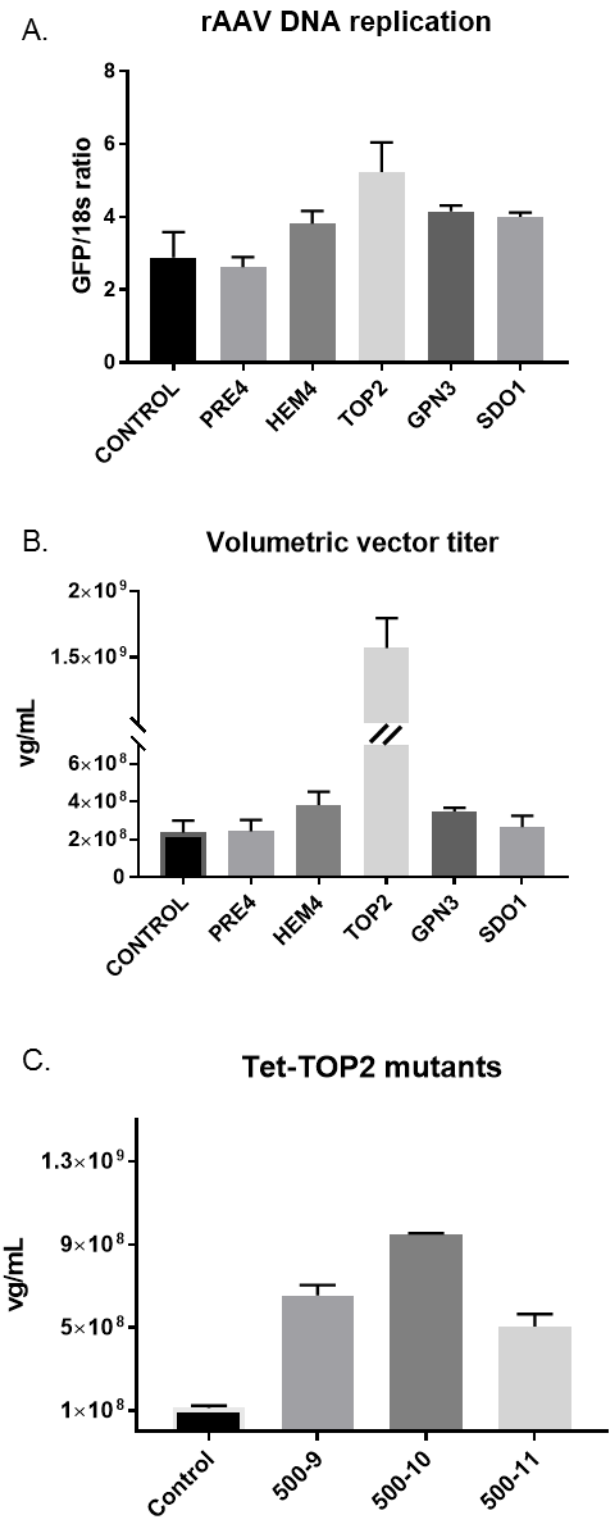
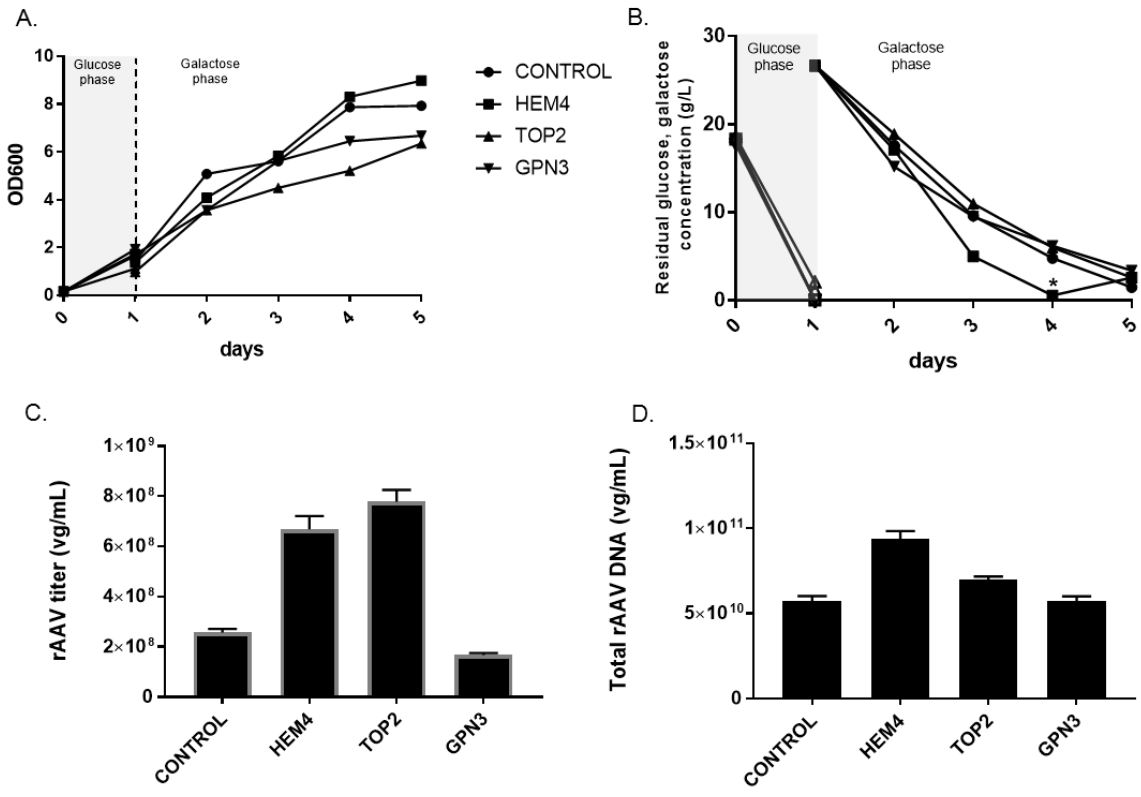


Fig 5. Evaluation of selected YPH500 overexpression strains. A pESC plasmid containing GAL10 promoter was used to express 5 protein candidates. Fermentation was performed

in 24 deep-well plates. In each case, the control represents YPH500 strain transformed with AAV 4-plasmid system. rAAV-GFP/18s rDNA ratio analysis (A) and benzonase-resistant rAAV DNA analysis (B) of cell lysate material are shown. Tet-TOP2 integrated yeast mutants were generated, and their performance was similarly assessed to complement plasmid overexpression data (C).



E.

	rAAV-GFP titer/OD (vg/mL)	rAAV-GFP titer/biomass (vg/mg DCW)	AAV DNA/biomass (vg/mg DCW)	Capsid:vg ratio
Control	3.24E+07	4.85E+07	1.08E+10	1414
HEM4	7.45E+07	1.23E+08	1.72E+10	905
TOP2	1.23E+08	1.87E+08	1.68E+10	767
GPN3	2.50E+07	3.40E+07	1.16E+10	3116

Fig 6. Cell growth analysis (A), glucose and galactose residual concentration (B), and vector productivity yields (C, D, E) of three YPH500 overexpression clones under controlled fermentation conditions. (*) Additional galactose feed was added to this condition at the indicated time.

Additional File 1. List of primers used in this study.

Name	Sequence (5'→3')
JA001	ATCTAAGAAACCATTATTATCATGACATTAACC
JA002	TCTTGAGATCCTTTTTTTCTGCGCGTAATCTGC
JA003	GATCAAAGGATCTGAGCGACCTCATGCTATACCTGAG
JA004	CGCGCAGAAAAAAAGGATCTCAAGACTTCGAGCGTCCCAAAACCT
JA005	ATGAGGTCGCTCAGATCCTTTGATCTTTTCTACGGGGT
JA006	GAATGGCGAATGGGATCCAATATCAAAGGAAATGATAGCATTGAAGG
JA007	TGATATTGGATCCCATTCGCCATTCAGGCTG
JA008	TCATGATAATAATGGTTTCTTAGATAAGAGCGCCCAATACGCAA
JA009	CGACTCACTATAGGGCCCATGTCAACTGAACCGGTAAGCG
JA010	CCATGTCGACGCCCTCAATCCTCTTCATTGAACGAAACATCTG
JA011	CGACTCACTATAGGGCCCATGTCTAGTCGTAAAAAAGTCAGAG
JA012	CCATGTCGACGCCCTTATTTATGGCGCTGGTATAATTCAATAG
JA013	CGACTCACTATAGGGCCCATGAATCACGATCCTTTCAGTTGG
JA014	CCATGTCGACGCCCCTAAATTTTTTGAGTACCGTAGCC
JA015	CGACTCACTATAGGGCCCATGCCTATCAATCAACCGTCG
JA016	CCATGTCGACGCCCTTAGTTATGCGTTGTATTATCTATGACTG
JA017	CGACTCACTATAGGGCCCATGTCTCGCGTTGGTGTC
JA018	CCATGTCGACGCCCCTATTCTTCGACATCTATTTGGTCG
DB307	ACCCTGAAGTTCATCTGCACCAC

DB308	CTGCCCGTGCCCTGGCCCAC
DB309	CCGTAGGTCAGGGTGGTCAC
DB438	GGACGTTTGGTTCTATTTTGTTGGTTTCTAG
DB439	CTGATGCCCCCGACCGTCCCT
DB440	TTCGCAGTAGTTAGTCTTCAATAAATCCAAG
DB633	CGCGTCTAGACTCCCGGAGCGAGGAAAACTCTTCAATAGTTTTGCCAG
DB634	CGGTTAATTAACTCCCGGAGTATGTCACCGACGCAAAGAGATGAAAC
DB668	GCCGGATCCGGGTTGAATTCGAATTTTCAAAAATTCTTAC
DB669	GCCGCGGCCGCGGGTTTTTTCTCCTTGACGTTAAAG
DB560	CGCCGAATTCGTACGCTGCAGGTCGACAAC
DB561	CGCGAATTCGTCGACGCATAGGCCACTAGTGGATCTG
DB574	CTGTTTCAGTTAAAGGAGTTTATAACGACCAGCACGGCTAACTCCCAG GTTCGGATGTTC
DB575	TTTCTGATATTTATCAGAGGCGCTTACCGGTTCAGTTGACATGCCCCCG AATTGATCCGG

CHAPTER 5.

Proteomic profiling of yeast host cell response to recombinant AAV2 vector production

ABSTRACT.

Recent studies on recombinant AAV vector production demonstrated the generation of infectious viral particles in *Saccharomyces cerevisiae*. Proof-of-concept results showed low vector yields that correlated with low AAV DNA encapsidation rates. In an attempt to understand the host cell response to rAAV production, we profiled proteomic changes throughout the fermentation process by mass spectrometry. By comparing a rAAV-producing yeast strain with a respective non-producer control, we identified a subset of host proteins with significantly different expression patterns during the rAAV induction period. Gene ontology enrichment and network interaction analyses highlighted five specific cellular events: protein folding, proteasomal degradation activity, oxidation-reduction processes, protein biosynthesis, and carbon metabolism. Specific fold change patterns of heat shock proteins and other stress protein markers suggested the occurrence of a cytosolic unfolded protein response during rAAV protein expression. A correlative increase in 20S-26S proteasome subunits and proteins involved in response to oxidative stress also suggested potential degradation of misfolded proteins, and cellular activities to ameliorate the effects of reactive oxygen species or other oxidants. In addition, several proteins involved in ribosome biogenesis, transcription and translation were downregulated; which are signs of generalized cellular stress. We tested the functional relevance of identified host proteins by overexpressing selected protein leads using low and high copy number plasmids. Increased vector yields were observed in SSA1, SSE1,

SSE2, CCP1, GTT1, RVB2, and GAL4 overexpression proteins. Overall, the obtained results underlined the utility of proteomic-based tools for the understanding and optimization of recombinant strains.

Keywords: Adeno-associated virus, vector production, *Saccharomyces*, mass spectrometry, protein profiling

BACKGROUND.

Recombinant adeno-associated viral vectors (rAAV) are emerging drugs for gene therapy applications. Ongoing animal and human trials are showing promising data for a variety of clinical conditions. The recent FDA approval of Luxturna (Spark Therapeutics) opens regulatory and commercialization pathways for new drugs, motivating other companies to pursue development of rAAV-based products.

Vector production at laboratory scale requires the interaction of several biological inputs (*e.g* plasmids, viral inoculum, auxiliary helper genes, and host cells) within a controlled cell culture environment (Aponte-Ubillus et al. 2018). This process is currently carried out using mammalian cells (HEK293, BHK) or insect cells (Sf9) genetically modified to express AAV proteins. Expression of the capsid proteins VP1, VP2, and VP3 and an assembly activating protein (AAP), leads to the formation of viral capsids in the nucleolus (Samulski and Muzyczka 2014). The expression of the non-structural proteins Rep68/78 and Rep52/40 triggers rAAV DNA replication and encapsidation of the generated single-stranded sequence (Balakrishnan and Jayandharan 2014). In mammalian cells, the expression of auxiliary adenovirus or herpesvirus proteins is necessary to complement rAAV production; the identified helper genes participate as trans-activating agents of AAV promoters, or modifiers of the host cell milieu (Geoffroy and Salvetti 2005).

As an alternative to complex eukaryotic production host cells, few groups have studied rAAV production in simpler organisms such as yeast. The potential development of a microbial platform would bring simple, lower-cost and scalable means for rAAV vector biomanufacturing. Proof-of -concept results showed very efficient rAAV DNA replication and capsid formation (Barajas et al. 2017; Galli et al. 2017). The low vector yields obtained in yeast could come as a result of an unbalanced expression of rAAV proteins, or a negative effect of host cell response to expression of foreign proteins and multiple forms

of rAAV DNA. In the present study, we aimed at increasing our understanding of yeast host cell response to rAAV expression by profiling proteomic changes throughout the production process. Bioinformatics analysis of change patterns helped us build secondary hypotheses regarding potential production constraints in the producer cell line. Finally, a protein overexpression strategy was performed to test these hypotheses, ultimately hoping to validate the appropriateness of our proteomics approach and improve vector specific productivity.

MATERIALS AND METHODS.

Strain and culture media

S. cerevisiae strain YPH501 (MATa/MATα ura3-52/ura3-52 lys2-801/lys2-801 ade2-101/ade2-101 trp1-Δ63/trp1-Δ63 his3-Δ200/his3-Δ200 leu2-Δ1/leu2-Δ1) was obtained from Agilent Technologies. 20% glycerol stocks were maintained at -80°C. YPD broth (1% yeast extract, 2% peptone, 2% dextrose) was used for culture start-up. Synthetic complete (SC) media lacking the appropriate amino acids was used for yeast transformation. SC media supplemented with 0.1M Na_2HPO4/NaH_2PO4 phosphate buffer and 2% glucose or 3% galactose was used for fermentation of rAAV-producing strains.

Plasmid design

Coding sequences for rAAV2 capsid and replication proteins were amplified from a pAAV RC2 plasmid and inserted into pESC 2-micron plasmids under the control of galactose-induced promoters, as described in Barajas et al. (Barajas et al. 2017). Briefly, all plasmids were generated using a pESC plasmid (Agilent Technologies) as vector. DB046 contains

a His3 selection marker and VP3 and AAP expression cassettes controlled by GAL1/10 bidirectional promoter. DB228 and DB138 plasmids contain a Leu2 selection marker and GAL-based Rep52 and VP1 expression cassettes. DB029 plasmid contains a Trp1 selection marker and GAL-based Rep78 and VP2 expression cassettes. JA001 plasmid consolidated the aforementioned AAV coding sequences into one plasmid. Plasmid DB040 is a pAAV-GFP-based plasmid (Cell Biolabs) containing Ura3 and 2-micron sequences. JA002 plasmid resembles DB040, with the difference that a Leucine marker was placed instead of the original Uracil marker. Protein overexpression plasmids were generated using DB327 (pESC(U)-2µ-GAL10), as a backbone vector. Specific primer sets (see additional file 1) were designed to amplify coding sequences from yeast genomic DNA. *Sma*I-digested DB327 plasmid and amplified sequences were ligated by Gibson assembly.

Proteomic profiling study design

The experimental design was aimed at identifying protein expression differences between rAAV-producing and non-producing YPH501 strains. The YPH501-rAAV strain was developed by transformation with plasmids DB046, DB138, DB029 and DB040. A control strain was generated by transformation with 4 pESC empty plasmids. Fermentation was performed in 250-mL shake flasks containing 50mL of SC medium + 2% glucose. Each strain was inoculated at an approximated cell density of 0.2 OD600nm. An orbital shaker was used to agitate flasks at a rate of 240 rpm. After 24 hours of culture, medium was exchanged with fresh SC medium + 0.1M phosphate buffer + 2% galactose to induce rAAV protein expression (Figure 1A). Culture was extended for four days after media exchange. Cell growth, pH, and AAV protein expression were monitored throughout the culture. This experiment was performed in triplicate, at three different times for both strains

(n = 9). Yeast biomass samples were taken on days 0, 2 and 3 post induction to analyze their proteomic profile.

Sample preparation

Sample were prepared for MS analysis following the methodology suggested by Paulo et al (Paulo et al. 2015). Cells were washed twice with water and suspended in a buffer containing 50mM HEPES (pH 8.5), 8M urea, 75mM NaCl, and protease inhibitors Complete Mini (EDTA-free) and PhosStop (Roche). Cell suspension was concentrated to approximately 1.5-2 OD600nm. Glass beads were added to the suspension in a cells: buffer: beads ratio of approximately 1:2:2. Cell suspensions were submitted to three homogenization cycles of 30 seconds each with 30-second rest intervals, using a Maxiprep 24 homogenizer. Lysates were centrifuged at 100g x 5min and supernatant was aliquoted for further treatment. Sample aliquots were reduced by incubation in 5mM tris 2-carboxyethyl phosphine (TCEP) for 25 minutes at room temperature. Alkylation was subsequently performed by 30-minute incubation with 10mM iodoacetamide. Immediately after, samples were incubated in 15mM Dithiothreitol (DTT), and protein fractions were separated by methanol chloroform precipitation. Protein concentration of all samples was monitored by using BCA assay kit (Thermo Fisher Scientific). Protein fractions were dissolved in 50mM HEPES + 0.05% Rapigest, and digested with trypsin (EMD Millipore) at 100:1 protein to protease ratio, for 6 hours at 37°C. Enzymatic digestion was stopped by addition of 1% formic acid. Prepared samples were flash frozen for further analysis.

Liquid chromatography - mass spectrometry

A chromatographic separation was performed using Acquity M-Class UPLC fitted with an Acquity HSS T3 column (1.8 µm, 1.0 × 150 mm, 100 Å; Waters Corporation). Peptides were separated with a reversed-phase gradient elution running 0.1% formic acid and 0.1% formic acid in acetonitrile (Burdick and Jackson) from 3% B to 40% B over 80 minutes at 25 µL/min. Prior to injection on the column, 2 µg peptide sample solutions were spiked with 500 fmol of Hi3 *E. coli* peptide internal standard mixture (Waters Corporation) for subsequent quantitation by the "Hi3" method. Proteomics data was acquired on a Synapt G2-Si mass spectrometer (Waters Corporation) operating in $HDMS^E$ mode. Raw mass spectrometry data was processed for proteomics analysis with Progenesis QI for Proteomics software (version 3.0, Nonlinear Dynamics). Chromatograms were aligned and normalized using the "all proteins" approach. Peptides were identified within Progenesis from a search of the Uniprot *Saccharomyces cerevisiae* database (UP000002311) appended to include key AAV protein amino acid sequences and Hi3 internal standard peptides. Peptide search criteria included 10 ppm mass measuring accuracy, fixed carbamidomethylation, variable methionine oxidation, and a 4% false discovery rate. Ion matching requirements were two or more fragments/peptide, three or more fragments/protein, and one or more peptides/protein.

Bioinformatics data analysis

Preliminary Progenesis protein profiling data was refined by establishing a cut-off confidence ID value of 15, and the presence of at least 2 unique peptides per hit. Principal component analysis was performed in JMP 12 software (SAS) to visualize sample clustering and to identify potential run outliers. The analysis was run under the default estimation method. Basic statistical analysis (mean, standard deviation, *p*-value) of MS data was performed using Progenesis QI software and Microsoft Excel. A cut-off *p*-value

of 0.05 was used to determine differential protein expression between conditions. Differentially expressed proteins over days 2 and 3 post-induction between rAAV-producing and control strains were identified. A new list was created with the accession ID of proteins whose fold change was higher than 25%. This accession ID list was inserted into DAVID bioinformatics database (https://david.ncifcrf.gov/home.jsp) for gene ontology enrichment analysis. Threshold *q*-value for this analysis was established at 0.015. Additional bioinformatics analysis was performed using STRING network interaction database (https://string-db.org/). A threshold false discovery rate value of 4% was set for gene ontology analysis, and a high confidence (0.7) interaction score for predicted interaction analysis.

Protein overexpression study

The effect of protein overexpression on vector yield was assessed. Results from protein profiling led to the identification of protein candidates that could potentially improve rAAV expression. Nineteen candidates were screened by using a rAAV 2-plasmid yeast system (JA001 and JA002) on which a third plasmid containing the coding sequence of a protein of interest controlled by GAL10 promoter was transformed. 4 clones per strain were isolated and grown in 24-deep well plates following the fermentation strategy described above. 500µL samples were taken on day 4 post induction, and benzonase-resistant vector yield from each yeast lysate was determined. Paired, one-tailed T-tests were performed between each variable and control strains to determine if the average yield of the variable strain was significantly higher.

rAAV analytical testing

Digital droplet PCR (ddPCR) was performed as described in Barajas (Barajas et al. 2017) to quantify benzonase-resistant rAAV DNA. Yeast treated material was diluted 100:1000-fold to target the ddPCR dynamic range. Five µL of diluted material were mixed with 20 µL of Taqman-based master mix (BioRad) including GFP primers DB307/DB309 and a FAM dye-labeled probe DB308. Droplets were generated by an automated droplet generator (Biorad), and amplified material was analyzed in a QX200 droplet reader (Biorad) using Quantasoft software (Biorad).

RESULTS.

To evaluate yeast proteome changes occurring as a result of rAAV expression, fermentation runs using a yeast rAAV-producing strain and a non-producing control were performed in triplicates at three different times. The results from fermentation runs are presented in figure 1B-1D. Most conditions showed consistent growth profile throughout the 5 days of fermentation. There was no significant difference in growth rate trends after galactose induction, suggesting that rAAV protein expression did not significantly impact cell growth. pH trends were also comparable between strains, showing a subsequent mild decrease over the 4 days of growth in galactose. rAAV protein expression was tracked in both strains, and mass spectrometry was used for their detection on day 3 post induction. Figure 1D shows the detection of Rep, Cap and AAP proteins only in the rAAV-producing strain and not in the control strain.

Treated samples from days 0, 2, and 3 post induction were processed for UPLC-MS analysis, as described in the methodology section. We performed principal component analysis (PCA) to visualize and corroborate clustering of proteomic data among yeast samples. We put more attention to yeast samples in which rAAV expression took place

(days 2 and 3 post induction). Analysis of all samples showed a small cluster formed by day 2 and 3 post induction, rAAV-producing strain samples; this variability was partially explained by transformed components 1 and 2 (Fig 2). An independent PCA analysis of day 2 and day 3 samples from both strains showed how transformed component 2 explained 24.8% of proteome sample variability. We can infer that the identity of the variables that conform component 2 might be implicated in significant biological processes associated to rAAV production.

We used Progenesis QI for mass spectrometry raw data processing. A total of 925 yeast proteins were identified, covering protein IDs present in several cellular structures such as cell wall, cytoplasm and nucleus. 670 proteins showed significant changes in protein accumulation during rAAV expression when compared to the control strain ($p < 0.05$). Out of this total, approximately 70% met the cut-off requirement for our profiling analysis (See supplemental file 2 for details). Quantification of differential gene expression hits among different biological samples showed a total of 234 hits with a fold change bigger than 25% between different strains during days 2 and 3 post induction (Fig. 3). We used the identity of this protein set to study host response' biological processes associated to rAAV protein expression. Gene ontology enrichment analysis was performed by using DAVID software. As seen in table 1, the analysis showed a total of 6 biological processes with a q-value lower than 0.05. Half of them are processes related to protein folding, whereas the others are related to oxidation-reduction processes and carbon metabolism. Further analysis performed with STRING software contributed to the generation of a prediction-based interaction network, based on the protein set ID and protein-protein interactions reported in the literature. Once the network was designed, we identified 5 important clusters of proteins that participate in biological processes that could directly or indirectly impact rAAV protein expression (Fig 4). These processes are protein folding/refolding, proteasome

activity, protein biosynthesis, response to oxidative stress, and carbon metabolism. After analyzing the potential implications of expression change of these proteins, we speculated that the over-expression of specific host cell proteins identified in the profiling could provide stress-relief activity to the yeast cell, and potentially contribute to vector yield improvement. We generated a set of low-copy and high-copy number plasmids containing expression cassettes for 19 yeast host cell proteins. We overexpressed these proteins in a yeast strain transformed with two rAAV plasmids (JA001 and JA002), using a total of 3 plasmids for each yeast strain variant. As shown in figure 5, expression of some of these proteins using high-copy number plasmids resulted in an improvement in rAAV vector yield. Variant strains SSA1, RVB2, SSE1, SSE2, CCP1, GTT1, and GAL4 showed significant increases in vector yield that go as high as 3-fold relative to the control strain. These results support the significance of the aforementioned biological processes and the functional relevance of selected host cell proteins on rAAV vector production when using *S. cerevisiae* as host.

DISCUSSION.

Time-course mass spectrometry analysis of yeast samples allowed us to survey changes in the proteomic profile of the rAAV-producing yeast strain. Bioinformatics and statistical tools played an important role in highlighting expression changes and clustering them as part of biological processes. With those results, we identified processes that significantly varied when rAAV2 proteins were expressed. We built secondary hypothesis that could potentially link these few biological processes to bottlenecks in rAAV vector production.

Our results from proteomic profiling highlighted events related to protein folding and conformational stress. Protein folding/refolding comprises cellular activities aimed at

shaping the native conformation of proteins (Gasser et al. 2008). In order to keep cellular homeostasis, cell responses are focused on correcting the conformation of misfolded proteins, either by refolding, sequestration, or degradation (Chen et al. 2011). Protein folding takes place in the endoplasmic reticulum (ER) and cytoplasm, and each compartment has its own arsenal of folding proteins capable of doing a variety of modification to the target protein. In recombinant strains, protein expression and processing machinery is mostly allocated in the ER, where the nascent protein is modified and prepared for secretion (Mattanovich et al. 2004; Gasser et al. 2008) . For this particular case, the rAAV expression cassettes encode 6 non-glycosylated viral proteins that undergo processing in the cytoplasmic compartment. Our proteomic profiling results highlighted overexpression of several cytoplasmic heat shock proteins (Hsp) and chaperones, during the last days of galactose induction. This correlation implies a link between protein folding changes and the expression of rAAV proteins. No differences in expression of foldases and other chaperones associated with ER processing were seen, which gives support to the hypothesis that saturation of the protein processing machinery might be taking place, potentially leading to protein misfolding at the cytoplasmic compartment.

Table 3 shows MS results regarding fold change activity of the principal heat shock proteins during induction time. Results showed a 2-fold increase in Kar2 protein (also known as BiP). This protein is a stress marker and its upregulation is usually linked to unfolded protein response (Hohenblum et al. 2004). This protein was one of the few ER-related proteins that changed in concentration after galactose induction of rAAV expression. The cytoplasmic proteins SSA1, SSA2, SSA4, SSE1 and SSE2 increased their expression levels more than 25%, and our protein overexpression studies confirmed their functional relevance for rAAV vector production. These chaperones belong to the

Hsp70 and Hsp110 families, and are implicated in protein folding activity at the cytoplasm (Bush and Meyer 1996; Dragovic et al. 2006). Big and small heat shock proteins showed different change in their expression patterns, which aligned with results reported by Geiler-Samerotte (Geiler-Samerotte et al. 2011). The authors referred to this particular phenomenon as cytoplasmic unfolded protein response, and it has been reported on other occasions when surface viral proteins are expressed in yeast (Čiplys et al. 2011). It is believed that chaperone action is crucial to mitigate negative impacts related to protein misfolding. Valaviciute et al (Valaviciute et al. 2016) evaluated the effect of overexpression and mild downregulation of Hsp90, Hsp70 and Hsp40 chaperones and co-chaperones during recombinant expression of VP1 hamster polyomavirus protein in yeast. Downregulation of cytosolic chaperones such as SSA1/SSA2, SSA3/SSA4, HSP82 and HSC82 had a negative effect on VP1-EGFP levels. In addition, mild overexpression of these proteins translated into a surplus of VP1 yield. Their results suggested that these sub group of proteins have a direct impact on protein processing, and by extension on active recombinant protein yield.

An additional function of the host cell proteins highlighted in our analysis is related to protein degradation. During the profiling analysis, changes in proteins that are components of the proteasome subunits such as RPN6, RPT7, and PRE7 were identified. Many cytoplasmic chaperones participate in ubiquitin-dependent and independent degradation processes, which include proteasome activity (Santamaria et al. 2003; Ben-Nissan and Sharon 2014). The increased expression of this particular set of proteins suggests that protein degradation activities took place during rAAV protein expression, likely by proteasome activity on misfolded rAAV proteins such as capsids. Empirical knowledge from western blot analysis showed VP capsid proteins of multiple sizes besides

the three expected sizes (data not shown), which would support the notion of potential protein degradation events occurring at the cytosol.

Among oxidation /reduction biological processes, we were interested to evaluate the host cell response to oxidative stress. Thioredoxins, catalases, superoxide dismutases, and glutathione transferases usually play an important role during oxidative stress (Gasch 2003). Several proteins with these functions were detected on the MS analysis, but we particularly focused on the ones that showed increased expression after induction. Gad1 and Gtt1 showed significantly increased protein levels after galactose induction. These two proteins participate in the metabolism of glutamate and glutathione, respectively; which indirectly impacts the intracellular redox potential and modulates the stress generated by toxic oxidants (Coleman et al. 2001; Grant 2001; Collinson and Grant 2003). Other overexpressed proteins were Ccp1, Gre3, and Ahp1. These proteins, present in cytosolic and mitochondrial compartments, are expressed under stress conditions, and are implicated in different metabolic routes that protect cells against oxidant damage (Lee et al. 1999; Charizanis et al. 1999; Aguilera and Prieto 2001). Increased expression of some of these antioxidant proteins, however, has been associated to diauxic shift in yeast strains. Since this metabolic event is common in strains that grow in glucose-galactose media transitions (Murphy et al. 2015), it is plausible to think that the specific increase of some of these proteins might have been triggered by metabolic events different from rAAV production. It is important to note that protein expression fold changes in our rAAV-producing strain were at least 25% higher than the respective changes seen in the non-producing, control strain, which would suggest that this specific protein expression fold change could have occurred in both strains but with stronger response in the rAAV-producing strain.

By looking at host proteins involved in protein biosynthesis, we noticed that several tRNA transferases, RNA polymerases and some elongation factors were downregulated, which suggest that a general suppression of translation as a result of cellular stress could have taken place. This phenomenon could also have been worsened by environmental causes such as the lack of nonfermentable carbon sources or another limiting nutrient in media, which tends to impact processes related to ribosome biogenesis and translational activity (Gasch 2003; Ashe and Bill 2011).

Finally, we generated a KEGG pathway to look at host proteins that participate in carbon metabolic routes (Fig 6). The main finding was that key proteins associated to glycolysis, tricarboxylic acid cycle, and oxidative phosphorylation were upregulated during the last days of galactose induction. These results are usually seen in yeast stains exposed to environmental stress (Gasch 2003) and suggest the need to increase metabolic fluxes around these pathways, to cope with major energetic constrains.

Based on our secondary hypotheses, we generated a protein overexpression strategy. We hypothesized that additional expression of host cell proteins would benefit stress-free cell metabolism. Results shown in our yeast model using low-copy and high-copy number plasmids supported our previous hypotheses and practical objective of a proteomics-guided optimization of the rAAV-producing yeast strain. In a few cases like with GAL4 overexpression, we saw that protein overexpression at low levels exerted a stronger yield improvement than expression using high copy number plasmids. In other cases such as YDJ1, protein overexpression using high copy number plasmids led to a vector yield lower than the one obtained with the control strain. These affirmations infer that modulation of chaperone/host protein levels is required to achieve optimal yields. Vector titer improvement was evidenced after overexpression of proteins related to protein folding, response to oxidative stress, and regulation of gene expression. In addition, we tested

GAL4 overexpression in the system. This approach has been explored by other groups that worked on GAL-controlled recombinant protein expression (Schultz et al. 1987; Sil et al. 2000; Stagoj et al. 2006). We believed it was appropriate to test in this strain because we suspected that the levels of this important molecular activator might be too low to promote optimal rAAV gene expression. Positive results were seen at low-copy and high-copy gene number, which suggested that GAL4 might be a limiting factor in the regulation of rAAV protein expression, and that its activity could potentially be combined with overexpression of the aforementioned proteins, ultimately giving synergistic improvement towards vector yield.

CONCLUSIONS

The present study provided a molecular snapshot of proteomic changes related to rAAV expression in yeast. Data analysis suggested that cytoplasmic unfolded protein response might be taking place exclusively in our rAAV-producing yeast strain, and secondary events such as proteasome degradation and oxidative stress might be associated to it. We used this information to look for potential molecular strategies for optimization, and a host protein overexpression strategy resulting in increased yields supported this effort on the yeast model. The presented results encourage us to continue the exploration of this novel rAAV-producing microbial system. Future studies would potentially be focused on similar approaches in more complex rAAV-producing eukaryotic systems, in an effort to improve specific vector productivity in industrial strains.

DECLARATIONS

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material: The datasets supporting the conclusions of this article are included within the article.

Acknowledgments: The authors would like to thank Mimi Roy, Harry Sterling, Yvette Tang, Tomas Cinek and Marc-Andre Robert at BioMarin Pharmaceutical Inc for their suggestions throughout the execution of the study.

REFERENCES.

1. Aguilera J, Prieto JA (2001) The Saccharomyces cerevisiae aldose reductase is implied in the metabolism of methylglyoxal in response to stress conditions. Curr Genet 39:273–283

2. Aponte-Ubillus JJ, Barajas D, Peltier J, Bardliving C, Shamlou P, Gold D (2018) Molecular design for recombinant adeno-associated virus (rAAV) vector production. Appl Microbiol Biotechnol 102:1045–1054 . doi: 10.1007/s00253-017-8670-1

3. Ashe MP, Bill RM (2011) Mapping the yeast host cell response to recombinant membrane protein production: relieving the biological bottlenecks. Biotechnol J 6:707–714 . doi: 10.1002/biot.201000333

4. Balakrishnan B, Jayandharan GR (2014) Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. Curr Gene Ther 14:86–100

5. Barajas D, Aponte-Ubillus JJ, Akeefe H, Cinek T, Peltier J, Gold D (2017) Generation of infectious recombinant Adeno-associated virus in Saccharomyces cerevisiae. PloS One 12:e0173010 . doi: 10.1371/journal.pone.0173010

6. Ben-Nissan G, Sharon M (2014) Regulating the 20S Proteasome Ubiquitin-Independent Degradation Pathway. Biomolecules 4:862–884 . doi: 10.3390/biom4030862

7. Bush GL, Meyer DI (1996) The refolding activity of the yeast heat shock proteins Ssa1 and Ssa2 defines their role in protein translocation. J Cell Biol 135:1229–1237

8. Charizanis C, Juhnke H, Krems B, Entian KD (1999) The mitochondrial cytochrome c peroxidase Ccp1 of Saccharomyces cerevisiae is involved in conveying an oxidative stress signal to the transcription factor Pos9 (Skn7). Mol Gen Genet MGG 262:437–447

9. Chen B, Retzlaff M, Roos T, Frydman J (2011) Cellular strategies of protein quality control. Cold Spring Harb Perspect Biol 3:a004374 . doi: 10.1101/cshperspect.a004374

10. Čiplys E, Samuel D, Juozapaitis M, Sasnauskas K, Slibinskas R (2011) Overexpression of human virus surface glycoprotein precursors induces cytosolic unfolded protein response in Saccharomyces cerevisiae. Microb Cell Factories 10:37 . doi: 10.1186/1475-2859-10-37

11. Coleman ST, Fang TK, Rovinsky SA, Turano FJ, Moye-Rowley WS (2001) Expression of a glutamate decarboxylase homologue is required for normal

oxidative stress tolerance in Saccharomyces cerevisiae. J Biol Chem 276:244–250 . doi: 10.1074/jbc.M007103200

12. Collinson EJ, Grant CM (2003) Role of yeast glutaredoxins as glutathione S-transferases. J Biol Chem 278:22492–22497 . doi: 10.1074/jbc.M301387200

13. Dragovic Z, Broadley SA, Shomura Y, Bracher A, Hartl FU (2006) Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of Hsp70s. EMBO J 25:2519–2528 . doi: 10.1038/sj.emboj.7601138

14. Galli A, Della Latta V, Bologna C, Pucciarelli D, Cipriani F, Backovic A, Cervelli T (2017) Strategies to optimize capsid protein expression and single-stranded DNA formation of adeno-associated virus in Saccharomyces cerevisiae. J Appl Microbiol 123:414–428 . doi: 10.1111/jam.13511

15. Gasch AP (2003) The environmental stress response: a common yeast response to diverse environmental stresses. In: Yeast Stress Responses. Springer, Berlin, Heidelberg, pp 11–70

16. Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E, Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A (2008) Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. Microb Cell Factories 7:11 . doi: 10.1186/1475-2859-7-11

17. Geiler-Samerotte KA, Dion MF, Budnik BA, Wang SM, Hartl DL, Drummond DA (2011) Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. Proc Natl Acad Sci U S A 108:680–685 . doi: 10.1073/pnas.1017570108

18. Geoffroy M-C, Salvetti A (2005) Helper functions required for wild type and recombinant adeno-associated virus growth. Curr Gene Ther 5:265–271

19. Grant CM (2001) Role of the glutathione/glutaredoxin and thioredoxin systems in yeast growth and response to stress conditions. Mol Microbiol 39:533–541

20. Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D (2004) Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris. Biotechnol Bioeng 85:367–375 . doi: 10.1002/bit.10904

21. Lee J, Spector D, Godon C, Labarre J, Toledano MB (1999) A new antioxidant with alkyl hydroperoxide defense properties in yeast. J Biol Chem 274:4537–4544

22. Mattanovich D, Gasser B, Hohenblum H, Sauer M (2004) Stress in recombinant protein producing yeasts. J Biotechnol 113:121–135 . doi: 10.1016/j.jbiotec.2004.04.035

23. Murphy JP, Stepanova E, Everley RA, Paulo JA, Gygi SP (2015) Comprehensive Temporal Protein Dynamics during the Diauxic Shift in Saccharomyces cerevisiae. Mol Cell Proteomics MCP 14:2454–2465 . doi: 10.1074/mcp.M114.045849

24. Paulo JA, O'Connell JD, Gaun A, Gygi SP (2015) Proteome-wide quantitative multiplexed profiling of protein expression: carbon-source dependency in Saccharomyces cerevisiae. Mol Biol Cell 26:4063–4074 . doi: 10.1091/mbc.E15-07-0499

25. Samulski RJ, Muzyczka N (2014) AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. Annu Rev Virol 1:427–451 . doi: 10.1146/annurev-virology-031413-085355

26. Santamaria PG, Finley D, Ballesta JPG, Remacha M (2003) Rpn6p, a proteasome subunit from Saccharomyces cerevisiae, is essential for the assembly and activity of the 26 S proteasome. J Biol Chem 278:6687–6695 . doi: 10.1074/jbc.M209420200

27. Schultz LD, Hofmann KJ, Mylin LM, Montgomery DL, Ellis RW, Hopper JE (1987) Regulated overproduction of the GAL4 gene product greatly increases expression from galactose-inducible promoters on multi-copy expression vectors in yeast. Gene 61:123–133

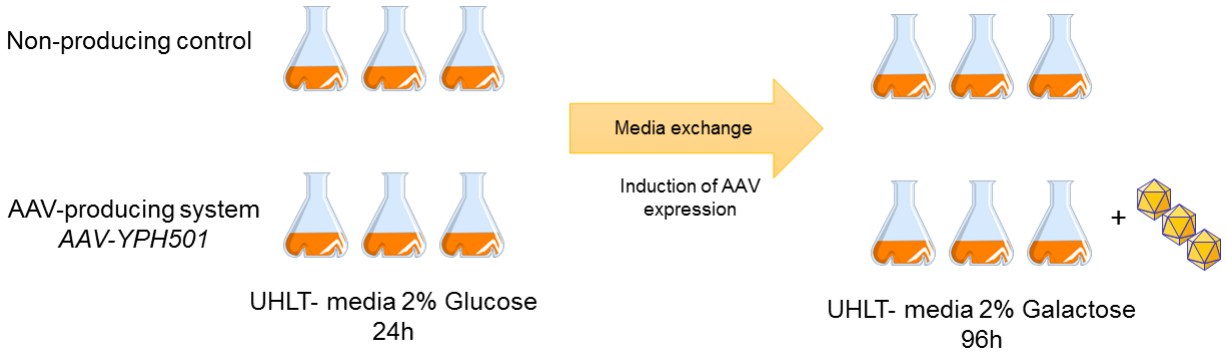
28. Sil AK, Xin P, Hopper JE (2000) Vectors allowing amplified expression of the Saccharomyces cerevisiae Gal3p-Gal80p-Gal4p transcription switch: applications to galactose-regulated high-level production of proteins. Protein Expr Purif 18:202–212 . doi: 10.1006/prep.1999.1177

29. Stagoj MN, Comino A, Komel R (2006) A novel GAL recombinant yeast strain for enhanced protein production. Biomol Eng 23:195–199 . doi: 10.1016/j.bioeng.2006.03.001

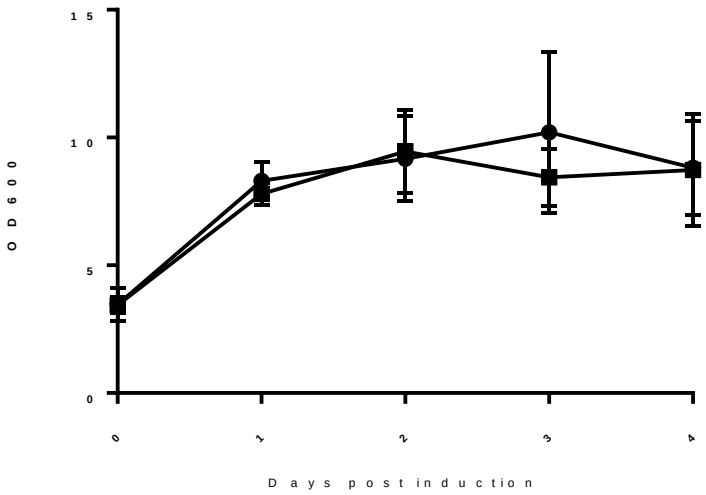
30. Valaviciute M, Norkiene M, Goda K, Slibinskas R, Gedvilaite A (2016) Survey of molecular chaperone requirement for the biosynthesis of hamster polyomavirus VP1 protein in Saccharomyces cerevisiae. Arch Virol 161:1807–1819 . doi: 10.1007/s00705-016-2846-3

FIGURES

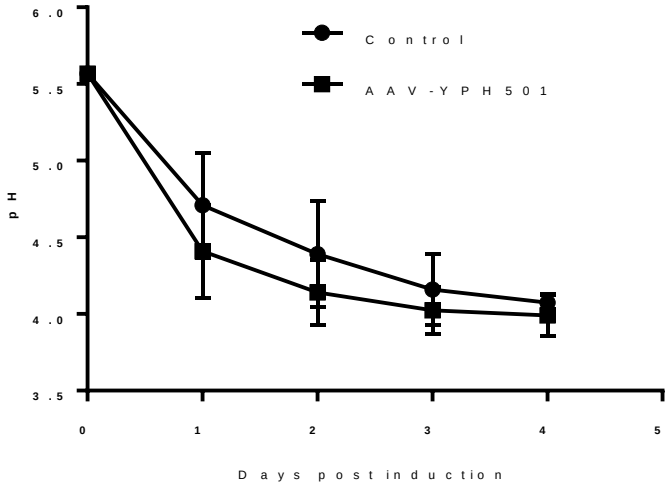
A.



B.



C.



D.

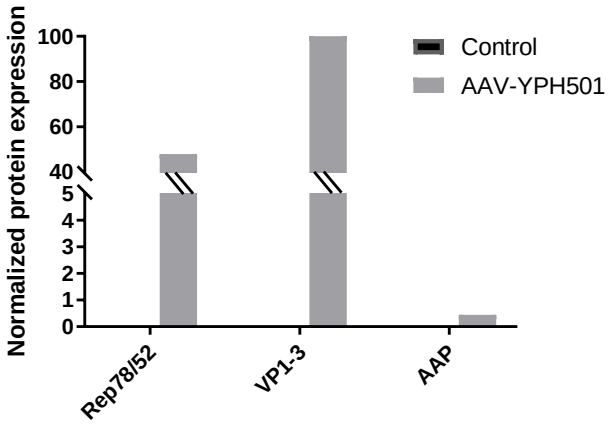
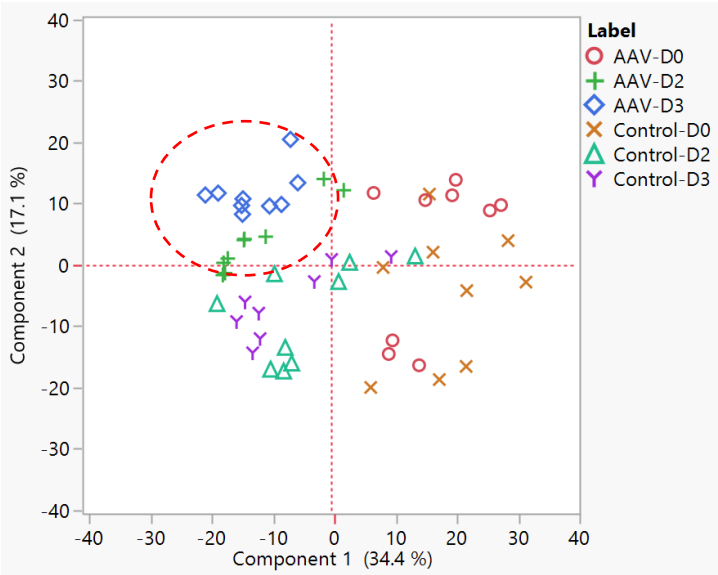


Fig 1. Proteomic profiling study. Outline of the experiment (A). Time-course yeast biomass (B) and pH monitoring (C) of all shake flasks conditions. Normalized rAAV protein expression in yeast samples on day 3 post galactose induction, detected by mass spectrometry (D).

A.



B.

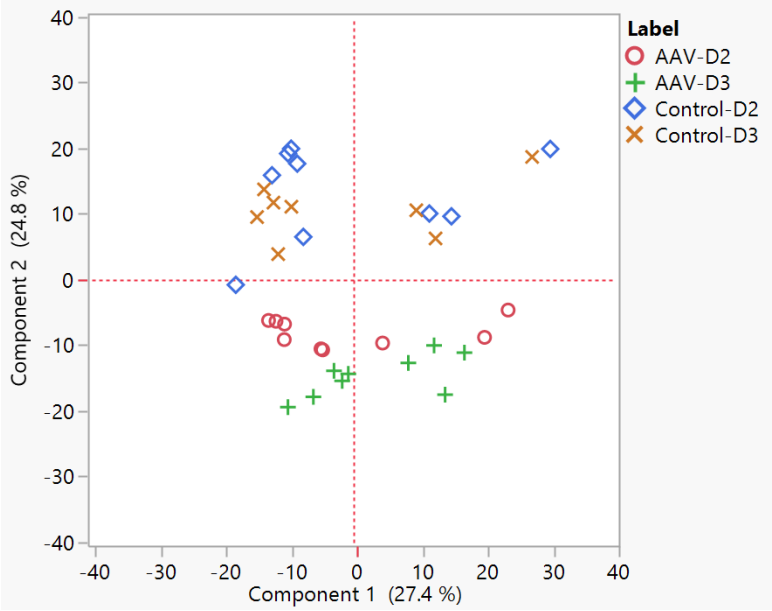


Fig 2. Principal component analysis. (A) Preliminary profile analysis of all yeast samples. (B) Profile analysis of all galactose-induced samples (AAV-D2, AAV-D3) and their respective controls (Control-D2, Control-D3).

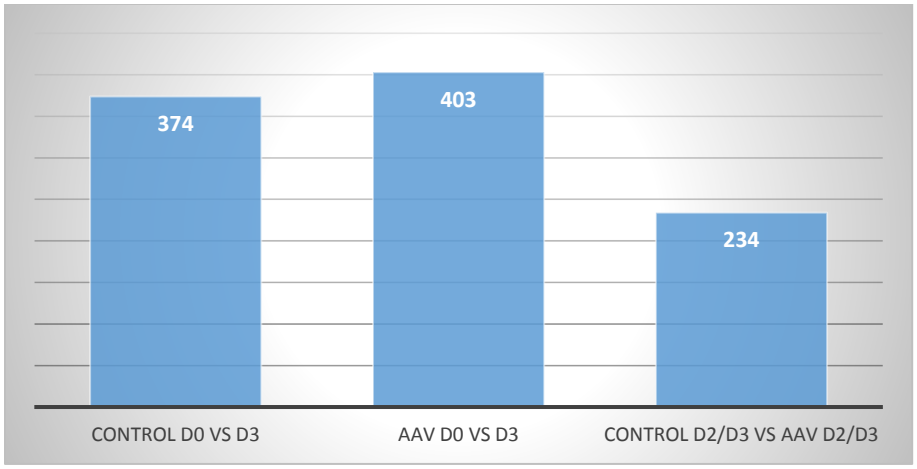


Fig 3. Quantification of differential protein expression (fold change > 1.25) among different sample sets.

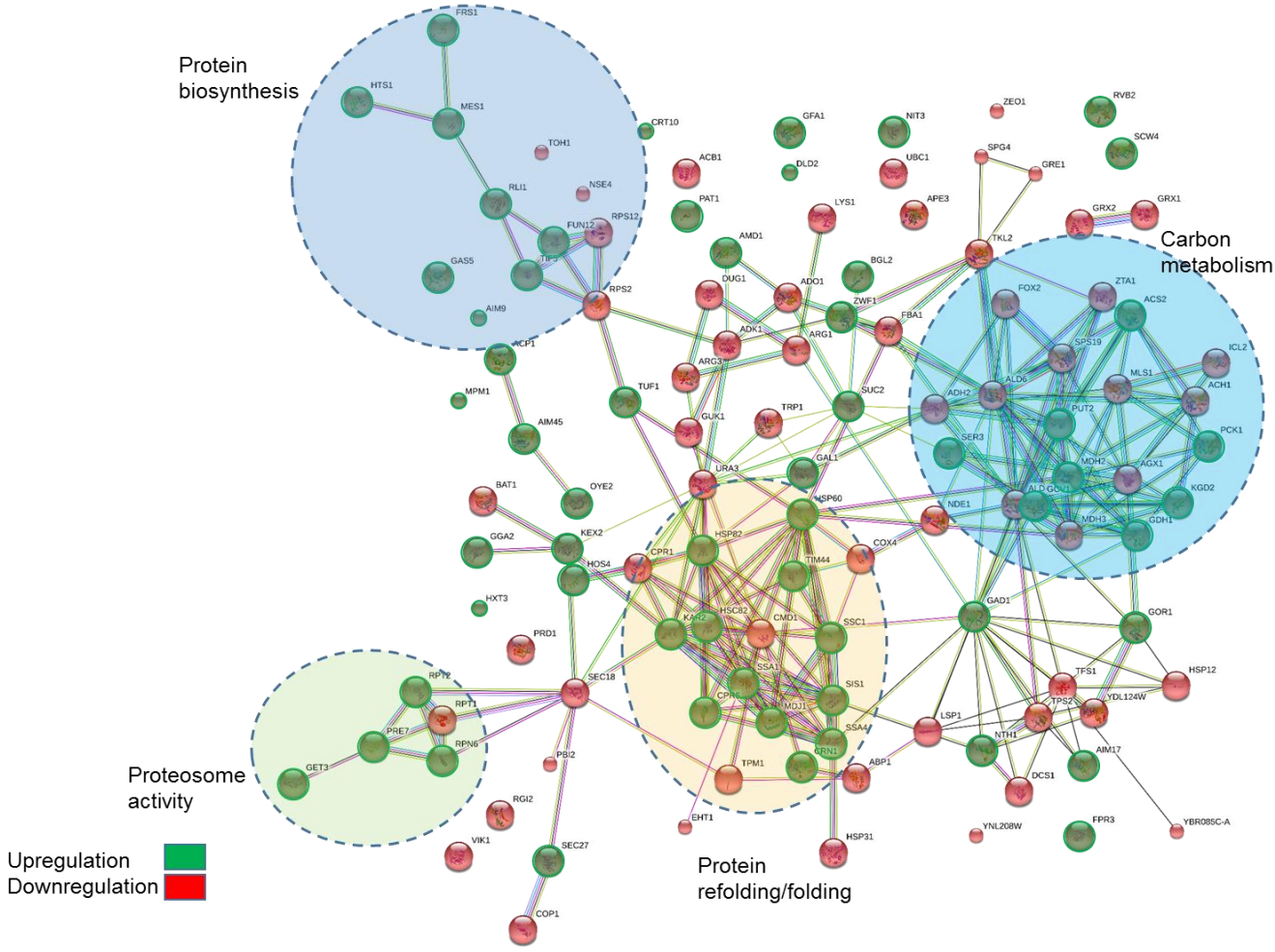
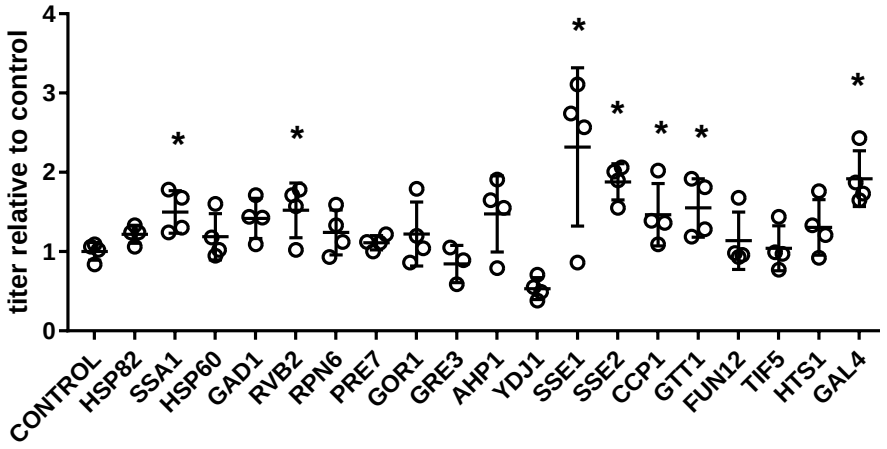


Fig 4. Predicted protein interaction network performed in STRING software. The analysis was done based on a list containing ID of proteins whose expression change was higher than 25%. Protein clusters that also appeared on the gene ontology analysis were identified and highlighted. Additional upregulation / downregulation information of each protein was included.

A.



B.

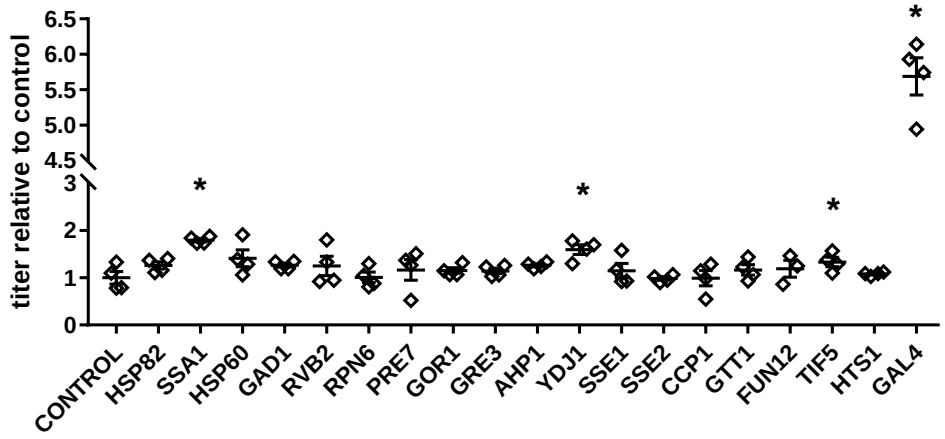


Fig 5. rAAV2 vector yield results from yeast protein overexpression strains. Nineteen strains were generated by transforming the control rAAV-producing strain with an additional 2-micron (A) or CEN (B) plasmid containing a GAL10-X host protein expression cassette. Benzonase-resistant vector yield results from each clone is presented as vector

titer relative to the control mean value. Bars represent mean and standard deviation (n=4). Asterisk represents conditions that are significantly higher than the control values, based on a paired, one-tailed t-test ($p<0.05$).

TABLES

Table 1. Gene ontology enrichment analysis performed using DAVID bioinformatics software. Table 1A consolidates enriched biological processes identified from proteomic changes in rAAV-producing yeast strain over time (Day 0 versus Day 3). Table 1B shows enriched biological processes identified from the differential profiling analysis (days 2-3) between rAAV-producing strain and the non-producing control strain.

A. AAV D0 vs D3

Term	Gene count	%	Q-value
Cytoplasmic translation	54	17.4	8.7E-25
Translation	70	22.6	4.4E-23
Metabolic process	46	14.8	4.2E-9
Glycolytic process	12	3.9	5.4E-6
Oxidation-reduction process	45	14.5	5.9E-6
Ribosomal small subunit assembly	11	3.5	4.6E-5
rRNA export from nucleus	9	2.9	1.6E-4
Gluconeogenesis	8	2.6	2.8E-3
Cellular amino acid biosynthetic process	17	5.5	3.5E-3
Protein refolding	7	2.3	7.4E-3
Carbohydrate metabolic process	16	5.2	1.8E-2
Protein folding	15	4.8	2.9E-2

B. Control D2-D3 vs AAV D2-D3

Term	Gene count	%	q-value
Oxidation-reduction process	25	20.2	4.5E-5
Protein refolding	7	5.6	8.9E-5
Metabolic process	19	15.3	2.4E-3
Protein folding	11	8.9	2.4E-3
'De novo' protein folding	4	3.2	1.2E-2
Carbohydrate metabolic process	10	8.1	1.5E-2

Table 2. Expression fold change of important chaperones and other host proteins implicated in cytosolic unfolded protein response and response to cellular stress

Protein name	Fold change	Function
SSA4	+2.04	Protein folding, cellular response to heat, SRP-dependent co-translational protein-membrane targeting and translocation
HSP82	+2.02	protein refolding, proteasome assembly, box C/D RNP assembly
HSC82	+2.02	Protein refolding, proteasome assembly box C/D RNP assembly
CPR6	+2.54	Protein folding, protein refolding
SSA1	+1.53	Translation, protein refolding, proteasome-mediated catabolic process, protein poly ubiquitination
YDJ1	+1.48	'de novo' protein folding, ER-associated ubiquitin dependent catabolic process, tRNA import into nucleus
SSC1	+1.63	Protein refolding, protein unfolding, protein import into mitochondrial matrix
HSP60	+1.57	Protein refolding, protein maturation, chaperone-mediated complex protein assembly
KAR2	+2.20	Unfolded protein binding, participates in ubiquitin-dependent ERAD pathway, protein import into the ER
AHP1	+1.06	Cell redox homeostasis, response to oxidative stress, response to metal ion
ZUO1	+1.02	Protein folding, regulation of translational fidelity, ribosomal subunit export from nucleus
PDI1	+1.03	Unfolded protein binding, protein disulfide isomerase activity
HSP12	-6.25	Lipid binding protein; involved in plasma membrane organization and response to oxidative, osmotic and heat stress

ADDITIONAL FILES

Table S1. List of primers sets used for amplification of sequences from yeast genomic DNA and further Gibson assembly of overexpression plasmids.

Name	Description	Sequence
JA043	HSP82-F	CGACTCACTATAGGGCCCATGGCTAGTGAAACTTTTG
JA044	HSP82-R	CCATGTCGACGCCCCTAATCTACCTCTTCCATTTCGG
JA045	HSC82-F	CGACTCACTATAGGGCCCATGGCTGGTGAAACTTTTG
JA046	HSC82-R	CCATGTCGACGCCCTTAATCAACTTCTTCCATCTCGG
JA047	CPR6-F	CGACTCACTATAGGGCCCATGACTAGACCTAAAACTTTTTTTG
JA048	CPR6-R	CCATGTCGACGCCCTCAGGAGAACATCTTCGAAAG
JA049	SSA1-F	CGACTCACTATAGGGCCCATGTCAAAAGCTGTCGG
JA050	SSA1-R	CCATGTCGACGCCCTTAATCAACTTCTTCAACGGTTGG
JA051	MDJ1-F	CGACTCACTATAGGGCCCATGGCTTTCCAACAAGGTG
JA052	MDJ1-R	CCATGTCGACGCCCTTAATTTTTTTTTGTCACCTTTGATC
JA053	SSC1-F	CGACTCACTATAGGGCCCATGCTTGCTGCTAAAAACATAC
JA054	SSC1-R	CCATGTCGACGCCCTTACTGCTTAGTTTCACCAGATTC
JA055	HSP60-F	CGACTCACTATAGGGCCCATGTTGAGATCATCCGTTG
JA056	HSP60-R	CCATGTCGACGCCCTTACATCATACCTGGCATTCC
JA057	GAD1-F	CGACTCACTATAGGGCCCATGTTACACAGGCACGGTTC
JA058	GAD1-R	CCATGTCGACGCCCTCAACATGTTCCTCTATAGTTTCTC
JA059	RLI1-F	CGACTCACTATAGGGCCCATGAGTGATAAAAACAGTCGTATC
JA060	RLI1-R	CCATGTCGACGCCCTTAAATACCGGTGTTATCCAAG
JA061	PAT1-F	CGACTCACTATAGGGCCCATGTCCTTCTTTGGGTTAG
JA062	PAT1-R	CCATGTCGACGCCCTTACTTTAGTTCTGATATTTCACCATC
JA063	RVB2-F	CGACTCACTATAGGGCCCATGTCGATTCAAACTAGTGATCC
JA064	RVB2-R	CCATGTCGACGCCCTTATTCCGTAGTATCCATGGCATC
JA065	RPT2-F	CGACTCACTATAGGGCCCATGGGACAAGGTGTATCATC

JA066	RPT2-R	CCATGTCGACGCCCTCACAAGTATAAACCTTCTAAATTTTCC
JA067	RPN6-F	CGACTCACTATAGGGCCCATGTCTCTGCCAGGTTCG
JA068	RPN6-R	CCATGTCGACGCCCCTAATACAAGACACTTGCCTTTTC
JA069	PRE7-F	CGACTCACTATAGGGCCCATGGCCACTATTGCATCAG
JA070	PRE7-R	CCATGTCGACGCCCTTAATCTCTTTTTAGCTCATAAAATTCTTTC
JA071	GOR1-F	CGACTCACTATAGGGCCCATGAGTAAGAAACCAATTGTTTTG
JA072	GOR1-R	CCATGTCGACGCCCTCAAACTAATGGCTTAGATTCATTGG
JA073	GRE3-F	CGACTCACTATAGGGCCCATGTCTTCACTGGTTACTCTTAATAAC
JA074	GRE3-R	CCATGTCGACGCCCTCAGGCAAAAGTGGGGAATTTAC
JA075	AHP1-F	CGACTCACTATAGGGCCCATGTCTGACTTAGTTAACAAGAAATTCC
JA076	AHP1-R	CCATGTCGACGCCCCTACAAATGAGCCAAGACAC
JA077	ZUO1-F	CGACTCACTATAGGGCCCATGTTTTCTTTACCTACCCTAACC
JA078	ZUO1-R	CCATGTCGACGCCCTCACACGAAGTAGGACAAC
JA079	YDJ1-F	CGACTCACTATAGGGCCCATGGTTAAAGAAACTAAGTTTTACG
JA080	YDJ1-R	CCATGTCGACGCCCTCATTGAGATGCACATTGAACAC
JA081	SSE1-F	CGACTCACTATAGGGCCCATGAGTACTCCATTTGGTTTAG
JA082	SSE1-R	CCATGTCGACGCCCTTAGTCCATGTCAACATCACC
JA083	SSE2-F	CGACTCACTATAGGGCCCATGAGCACTCCATTTGGC
JA084	SSE2-R	CCATGTCGACGCCCTTAATCAAGGTCCATGTTTTCATC
JA085	HSP10-F	CGACTCACTATAGGGCCCATGTCCACCCTTTTGAAGTC
JA086	HSP10-R	CCATGTCGACGCCCTTAGTCCTTGGCAATCTTAGCC
JA087	CCP1-F	CGACTCACTATAGGGCCCATGACTACTGCTGTTAGGC
JA088	CCP1-R	CCATGTCGACGCCCCTATAAACCTTGTTCCTCTAAAGTC
JA089	GTT1-F	CGACTCACTATAGGGCCCATGTCGTTGCCAATTATCAAAG
JA090	GTT1-R	CCATGTCGACGCCCTTAGAAATTGCTACCTAAAGCACG
JA091	SSA2-F	CGACTCACTATAGGGCCCATGTCTAAAGCTGTCGGTATTG
JA092	SSA2-R	CCATGTCGACGCCCTTAATCAACTTCTTCGACAGTTGG
JA093	HSP104-F	CGACTCACTATAGGGCCCATGAACGACCAAACGCAATTTAC
JA094	HSP104-R	CCATGTCGACGCCCTTAATCTAGGTCATCATCAATTTCCATAC
JA095	FUN12-F	CGACTCACTATAGGGCCCATGGCGAAAAAGAGTAAAAAGAACC

JA096	FUN12-R	CCATGTCGACGCCCTCATTCGATGCCGAAAACGACC
JA097	TIF5-F	CGACTCACTATAGGGCCCATGTCTATTAATATTTGTAGAGATAATCATG
JA098	TIF5-R	CCATGTCGACGCCCCTATTCGTCGTCTTCTTCATCATC
JA099	HTS1-F	CGACTCACTATAGGGCCCATGCTTAGTAGATCACTAAATAAAGTAG
JA100	HTS1-R	CCATGTCGACGCCCTTATAATCCTTTAATTAAACGAGTGACC
JA101	GAL4-F	CGACTCACTATAGGGCCCATGAAGCTACTGTCTTCTATCG
JA102	GAL4-R	CCATGTCGACGCCCTTACTCTTTTTTTGGGTTTGGTGG

CHAPTER 6.

Bioprocess development studies on recombinant adeno-associated virus 2 (rAAV2) vector production using *Saccharomyces cerevisiae*

ABSTRACT.

Recombinant AAV vector production in simple eukaryotic organisms such as yeast has been explored by a handful of research groups. Preliminary data demonstrated generation of infectious rAAV2 vector particles in *S. cerevisiae* at very low volumetric yields. Molecular and process optimization studies are necessary before assessing the real potential utility of this platform in industrial settings. In the present study, we evaluated how rAAV plasmid design and bioprocessing parameters impact vector yield in yeast.

The rational molecular design was conducted initially by comparing transient versus stable rAAV2 expression. Vector yield in recombinant yeast when using 2-micron plasmids far exceeded the performance of stable clones which have rAAV cassettes integrated in the yeast chromosome, suggesting the importance of DNA copy number in AAV protein expression. We also evaluated multiple arrangements of rAAV expression cassettes consolidated into 2 or 4 plasmids. Vector yield was 5-fold higher when separated into 4 plasmids. Subtle differences were noticed when different auxotrophic markers were used for plasmid selection, suggesting the importance of balancing AAV gene copy number to prevent negative cellular effects as a result of protein overexpression. Evaluation of different codon-optimized cassettes suggested that high levels of Rep78 can be detrimental for rAAV2 production.

Yeast fermentation process was characterized, taking into account the use of galactose as induction media component for rAAV protein expression. Time-course analysis of rAAV DNA replication and vector yield showed day 4 post induction to be optimal for cell harvest. Shake flask experiments were designed to screen for optimal pH, induction temperature and media formulation for production of rAAV2-GFP. 250ml volume stirred-tank reactors were used to define a growing strategy based on a preliminary growth phase using glucose exponential feeding, followed by multiple galactose bolus feed additions. Evaluation of optimal bioreactor parameters (pH, induction temperature, DO %) was carried out based on design of experiments (DoE) response surface methodology, using rAAV vector yield as primary output. Combination of shake flask studies and DoE experimental results led to a cumulative increase in vector yield as high as $1.5x10^{10}$ vg/ml. This value represents approximately a 50-fold increase of the initial yields reported in the preliminary proof-of-concept study (Barajas *et al*. 2017). Provided that vector product maintain acceptable quality attributes, this novel system shows a lot of potential for cost-effective and scalable production of rAAV2 vectors.

Keywords: Adeno-associated virus, vector production, *Saccharomyces*, fermentation

BACKGROUND.

Adeno-associated viruses (AAV) are single-stranded DNA viruses classified as *Dependoparvovirus* (Daya and Berns 2008). The potential utility of AAV as viral vector was discovered in the mid 80's, when it was demonstrated that this virus could be genetically modified to transfer specific genes into mammalian cells (Tratschin et al. 1984; Tratschin et al. 1985). The use of recombinant AAV gained more importance after several pre-clinical and clinical studies evidenced its non-toxic profile and potential long-term transgene expression. Overall, this genetic therapy is expected to become a safe, one-dose solution for multiple genetic deficiencies.

Production of recombinant AAV is traditionally carried out using mammalian cells such as HEK293 or BHK. The upstream manufacturing process requires the insertion of AAV genes into the host cell via plasmid transfection or recombinant viral transduction. The expression of the required structural (VP1, VP2, VP3) and non-structural (Rep78/68, Rep52/40, assembly-activating protein) proteins is driven by p5, p19 and p40 native promoters (Aponte-Ubillus et al. 2017). The gene of interest, flanked by AAV inverted terminal repeats (ITRs), is provided on a separate construct. ITRs contain AAV-specific recognition sequences required for DNA replication and encapsidation (Samulski and Muzyczka 1999). Furthermore, the addition of adenovirus or herpesvirus auxiliary helper genes is required to promote optimal AAV vector production. Insect cells are also used to produce recombinant AAV. This alternative platform normally uses 2 or 3 recombinant baculovirus strains to insert the required AAV genes and transgene of interest into the host. Baculovirus' promoters such as ΔIE10 and polh drive the expression of VP and Rep proteins (Urabe et al. 2002). Specific vector productivity on these two platforms range between 10^3 to 10^5 vector genomes (vg) per cell, which in suspension-adapted cultures usually translates into a volumetric yield of 10^9 to 10^{11} vg/mL.

Few groups have reported the production of recombinant AAV in simpler eukaryotic cells such as yeast (Backovic et al. 2012; Barajas et al. 2017; Galli et al. 2017). A potential AAV-producing microbial system would represent a more affordable and scalable platform to generate AAV vectors, provided that the final product maintain acceptable product quality attributes. Barajas *et al* (Barajas et al. 2017) showed formation of AAV vector particles at a volumetric yield of approximately $2x10^8$ vg/mL. Results suggested that DNA encapsidation is the limiting step in the process, since total rAAV DNA and total capsid yields exceeded the previously mentioned yield by 10^2-10^3 fold. Molecular and process optimization studies could lead to the improvement of volumetric yields. In the present work, we aim at developing an optimized rAAV vector production process in *Saccharomyces cerevisiae*. We evaluated how rAAV2 molecular plasmid design and bioprocessing conditions impact volumetric vector yield. We used shake flask and bioreactor data to design a 2-stage, fed-batch fermentation process for production of rAAV2-GFP vector particles.

MATERIALS AND METHODS.

Strains and culture conditions

S. cerevisiae strain YPH501 (MATa/MATα ura3-52/ura3-52 lys2-801/lys2-801 ade2-101/ade2-101 trp1-Δ63/trp1-Δ63 his3-Δ200/his3-Δ200 leu2-Δ1/leu2-Δ1) was obtained from Agilent Technologies. 20% glycerol stocks were maintained at -80°C. YPD broth (1% yeast extract, 2% peptone, 2% dextrose) was used for culture start-up. AAV-YPH501 strain was used for most optimization studies. This strain was generated by transforming YPH501 strain with four plasmids (DB029, DB040, DB155, DB149) that contained all rAAV2 expression cassettes plus the ITR-flanked GFP as described in Barajas *et al*

(2017). Different plasmid designs were tested as mentioned later, using YPH501 strain as a basis.

Deep-well plate fermentation was performed to screen clones with different rAAV plasmid configurations. Yeast clones were inoculated into 24-deep well plates containing synthetic complete (SC) media lacking the appropriate amino acids and supplemented with 0.1M Na_2HPO_4/NaH_2PO_4 phosphate buffer and 2% glucose. A temperature-controlled plate shaker (VWR) was used to incubate the cells at 29±1°C and 550 rpm agitation. After approximately 48 hours, the cultures were washed and fresh SC medium + 0.1M Na_2HPO_4/NaH_2PO_4 phosphate buffer + 3% galactose was provided. The fermentation was extended for 4 days, maintaining the same operating conditions. Shake flask fermentation was carried out to characterize yeast growth and productivity profile, and to screen the effect of media components, temperature and pH in vector productivity. 125-ml shake flasks were incubated in a MaxQ orbital shaker (Thermo Scientific) at 30°C temperature and 240rpm of agitation. Media composition was identical to the one used for deep-well fermentation.

Bioreactor fermentation was performed by using a Dasbox system with single-use Bioblu 0.3f fermenters (Eppendorf). For yeast culture at densities lower than an OD600nm of 15, a yeast starter culture was used to inoculate a 200mL working volume bioreactor, to a target OD600nm of approximately 0.3. Cells grew in SC medium with 2% glucose for over 24 hours, and a 20X concentrated SC-medium with 3% galactose was added immediately after. The fermentation was extended for 96 hours and then yeast cells were harvested. For yeast culture at densities higher than OD600nm of 15, cells were inoculated into 250ml stirred-tank vessels containing 100mL of a modified SC medium composed of 2g/L glucose, 6.7g/L yeast nitrogen base (YNB) without amino acids (Sigma), 5.6g/L amino acid dropout mix (Sigma), 5.8g/L YNB without amino acids and ammonium sulfate

(Sigma), 3g/L Na_2HPO_4; 15g/L NaH_2PO_4, and 100µg/mL carbenicillin. Target inoculation OD600nm was approximately 1.5. Higher densities were also explored (3 and 6 OD600nm) to evaluate their effect in cell growth and vector productivity. After an hour of batch growth in modified SC medium, a glucose feed was provided using an exponential feeding rate, targeting growth rate ranges between of 0.13-0.18 h^{-1}. Alternative glucose feeds were evaluated during this study. Feed A contained the following components: 500g/L glucose, 30 g/L YNB, 10g/L amino acid dropout mix, 40g/L YNB without amino acids and ammonium sulfate, and 40g/L Hy-case SF. Glucose feed (SC-2) contained the following components: 500g/L glucose, 30 g/L YNB, 70g/L amino acid dropout mix, and 70g/L YNB without amino acids and ammonium sulfate. Glucose feed SC-1 has similar composition than SC-2, with the difference that it contains 5g/L Hy-case SF. Once the target cell density was reached, exponential glucose feeding was stopped and a galactose feed was supplemented as a bolus, to induce rAAV protein expression. The galactose feed was composed of 400g/L galactose, 60g/L YNB, 28g/L amino acid dropout mix, and 9g/L lysine hydrochloride. This feed was provided twice a day for the following 4 days until termination. The standard bioreactor operating parameters were 550 rpm agitation, 30°C temperature, 4.8 pH and dissolved oxygen (DO %) set point of 30%. A gas control loop controlling impeller agitation, oxygen flow rate and total gas flow rate was set to maintain DO % present values close to the defined set point. A pH control loop was also set, and used 2M KOH to neutralize the medium as needed. An off-gas analyzer was used to monitor oxygen transfer rate, carbon dioxide transfer rate, and respiratory quotient (RQ) during the fermentation runs.

Recombinant plasmids

A number of plasmids were designed by inserting galactose-inducible expression cassettes for rAAV2 capsid and replication proteins into pESC 2-micron plasmids (Table 1) as described previously in Barajas *et al* (Barajas et al. 2017). VP1, VP2, VP3, Rep78, Rep52, and AAP expression cassettes were assembled into 1 or 3 plasmids. An additional assembly containing all ALL1 cassettes was integrated into the host genome. Another group of plasmids was designed to carry the transgene of interest flanked by AAV2 inverted terminal repeats (ITRs). To generate them, a yeast 2-micron sequence, an auxotrophic marker and GFP transgene sequence were inserted into a pAAV plasmid.

Media screening experimental design

A definitive screening experiment was designed with the objective to identify media components that exert either positive or negative effect on rAAV vector yield (vg/mL culture) or vector productivity (vg/OD600nm). JMP 12.1 software (SAS) was used to define experimental settings. Forty seven media components were screened at two levels for main effects and second-degree interactions. A total of 97 runs, including some center points, were divided into 4 blocks (Table S1).

Three 1L shake flasks containing 200mL yeast cultures were started in SC-glucose medium and incubated for 24 hours at standard operating conditions. After that period, the culture volume was consolidated and later split in twenty four, 25mL aliquots. The exhausted medium of each aliquot was exchanged with SC-Galactose fresh medium and transfer into a 125ml shake flask. Right after, a 10% fixed-volume feed based on variable formulations of the screened components was added to their respective shake flask. The runs were extended for four days until termination. Samples were taken on days 3 and 4 post galactose induction to analyze the two aforementioned performance outputs.

Bioreactor optimization experimental design

A response surface design was proposed to identify the optimal setpoint values for pH, DO % and temperature. Three, 2-level continuous variables and a blocking factor were considered for the design. By definition, this study evaluated main effects, second-degree interactions and quadratic effects. The operational features for each bioreactor were similar to the ones previously described in the methodology, using different setpoints for the 3 variables from the start of each fermentation run. The study included a total of 14 runs divided in 2 blocks (Table 2). Samples were taken during days 3 and 4 post galactose induction. The analyzed output was volumetric rAAV2 vector yield (vg titer/mL).

rAAV analysis and analytical testing

Digital droplet PCR was performed as described in Barajas *et al* (Barajas et al. 2017) to quantify benzonase-resistant rAAV DNA. Yeast treated material was diluted 100-1000 fold to target the ddPCR dynamic range. Five µL of diluted material were mixed with 20 µL of Taqman-based master mix (BioRad) including GFP primers (DB307/DB309) and a FAM dye-labeled probe (DB308). Droplets were generated by an automated droplet generator (Biorad), and amplified material was analyzed in a QX200 droplet reader (Biorad) using Quantasoft software (Biorad).

To determine to percentage of plasmid-containing cells, yeast samples were taken at time of galactose induction and cultured on YPD + 2% glucose, and SC medium + 2% glucose lacking the appropriate amino acids. The percentage was determined by dividing the colony forming unit (cfu) in SC medium by cfu in YPD medium and multiplying it by 100.

Western blot was used to determine the expression of rAAV VP1,2,3, Rep52 and Rep78 proteins as described in Barajas et al (2017). 1/1000 B1 anti-VP antibody (American Research Products) and 1/300 303.9 anti-Rep antibody (American Research Products) were the secondary antibodies used for specific protein detection.

RESULTS.

Evaluation of different plasmid arrangements

rAAV protein expression was evaluated in yeast cells transformed with different 4-plasmid and 2-plasmid arrangements. Rep/Cap western blot results showed attenuated expression levels in all 2-plasmid constructs, especially in VP1 and VP3 expression (Fig 1b). Additional analysis of total rAAV DNA per 18s DNA ratio evidenced higher ratios in all 4-plasmid arrangements, relative to their respective 2-plasmid constructs (Fig 1c). In an effort to improve rAAV DNA replication, additional codon-optimized sequences were evaluated (ALL2 and ALL4). However, PCR analysis showed comparable results between Rep78-optimized and non-optimized constructs. No significant benefits in rAAV DNA replication were observed in ALL2/ALL3/ALL4 constructs. Therefore, ALL1 original construct was used for further experiments.

In figure 2a, vector yield was compared among 4-plasmid and 2-plasmid system clones. Results showed 5-fold higher rAAV vector yield of the 4-plasmid system, relative to 2-plasmid variants with different auxotrophic marker configuration. The percentage of plasmid-containing cells was analyzed at galactose induction time. Results seen in figure 2b showed comparable values, suggesting that at small scale fermentation, plasmid stability is low (30-50%) but comparable among all constructs. Among 2-plasmid constructs, we saw small but significant differences in vector yield, which suggest that the

nature of the auxotrophic marker might be influencing plasmid number, and expression levels by extension. We decided to compare these results against a strain that contains all AAV expression cassettes integrated in its genome. Stable rAAV expression showed 100-fold lower titers compared to the 4-plasmid system, suggesting that copy number for stable production is a limiting factor for vector yield. A final evaluation of codon-optimized sequences was done only with a 4-plasmid construct and results are shown in figure 2c. Vector yield results evidenced that AAP and Rep52 optimized sequences must be present in the plasmid construct to obtain improved yields, as seen in our previous study (Barajas et al. 2017). Moreover, we observed an impairment in vector productivity when Rep78 sequence was codon-optimized.

Preliminary production studies in shake flasks

The yeast fermentation production process in shake flasks was defined as batch with medium replacement at induction time. We were initially interested in identifying the optimal length for the process, reason why we did a time-course analysis of rAAV DNA replication and vector yield. ddPCR results in figure 3a showed the start of rAAV DNA replication at 24 hours post induction, reaching peak levels at 48 hours post induction. Moreover, detectable levels of benzonase-resistant rAAV particles were seen at 48 hours post induction, with peak levels reaching at 96 hours post induction (Fig 3b). Based on the previous results, we decided to extend the duration of the process to a total of 5 days, which included 1 day of growth in medium with glucose and 4 days of growth in fresh medium with galactose. One-factor-at-a-time studies were performed to investigate the main effect of temperature and pH on specific vector productivity. Preliminary results suggested a strong effect of temperature on rAAV vector yield, especially at temperatures higher than 30°C, which showed a plunge in productivity as low as 20-fold (Fig 3c). In

order to evaluate the effect of pH in shake flasks, different buffer formulations were used to evaluate different pH ranges. Comparable productivity was seen in cultures with starting pH ranging from 4 to 5.8. A 2.5-fold decrease in productivity was seen in a conditions with a starting pH of 6.8 (Fig 3d, 3e). This information was important to design a basic bioreactor fermentation process.

A media optimization study was performed in shake flasks to identify components that could boost rAAV vector yield. A DoE definitive screening design was used to evaluate main effects and interactions of 48 media components. Results shown in figure 4 exhibited a wide range of values for volumetric vector yield (vg/mL) and specific vector productivity (vg/OD). We continued with further analysis using vector yield as main output. Data analysis in JMP software revealed lysine exerted the strongest positive effect, whereas histidine and zinc sulfate showed a small but significant negative effect. To confirm the data from the model generated by JMP software, 4 experimental conditions were run in parallel to the control. Conditions 46, 95 and one condition suggested by the maximum desirability profiler showed 2-3 fold improvement in vector yield. We also included one condition with lysine and 4 top components from the model with high estimate but not significant effect (pyridoxin, cysteine, ferric chloride and myo-inositol). This condition showed comparable results to the top conditions (~$1x10^{10}$ vg/mL). This information was used for future media replacement and bioreactor fed-batch runs.

Fermentation studies in 250mL bioreactors

The first objective was to develop a fed-batch fermentation process based on the information generated from the medium-replacement strategy performed in shake flasks. We used most of shake flask operating parameters to perform a fermentation run in a

bioreactor. Four bioreactors were inoculated at an OD600nm density of 0.3. Cells grew for approximately 24 hours in glucose-enriched media. Constant glucose monitoring from the fermenter medium helped us determine the optimal time for induction when glucose levels dropped below 0.5g/L). A 20X galactose feed using the components studied in the previous media optimization design, was added to the tank. A satellite run was done in a shake flask. Four different galactose induction strategies were evaluated: single 20X bolus feed, 20X continuous feed, 20X continuous feed + glucose, and 20X continuous feed + raffinose. As seen in figure 5, a comparable growth profile is evidenced among conditions except the one with 20X continuous feed, which showed lower growth. Vector yield results suggested that the bolus feed is the most effective galactose feed strategy. Interestingly, the satellite run performed in shake flask showed 5-fold lower vector yield. With the objective to intensify the fermentation process, we decided to use these data to evaluate an alternative fed-batch strategy that promoted exponential growth to moderate cell densities that would later be induced by multiple additions of galactose feed.

The next experiment evaluated the effect of glucose feed and exponential feeding rate on rAAV vector yield. Based on preliminary attempts, we identified 3 feed formulations (Feed A, SC-1, SC-2) that supported cell growth to moderate levels (data not shown). We wanted to investigate if the exponential feeding rate, which modulates yeast growth rate in culture, could impact vector yield. Results in figure 6 showed that the high feeding rate, which targeted a cell growth rate of $0.18h^{-1}$, raised the yeast metabolic activity during growth, as evidenced by the carbon transfer rate and respiratory quotient values obtained from the off-gas analyzer. Cultures fed at higher rates showed respiratory quotient values higher than 1.2, which correlated with an increase of fermentative metabolism as evidenced by high ethanol production (>15g/L) and high DO % in the fermenters, relative to the conditions that were fed at low rates (target cell growth rate of $0.13h^{-1}$). Analysis of

residual galactose in culture during the last 4 days of fermentation showed that the conditions fed at high rate also consumed less galactose. Yeast biomass results showed lower biomass profiles, suggesting that feeding rate does impact yeast growth. Vector yield results evidenced a lower productivity on the aforementioned conditions, suggesting that high glucose feeding rates have a direct or indirect negative impact on vector yield. Out of the three feeds used for yeast growth, SC-2 promoted better growth and vector yield. SC-2 is the yeast extract-free formulated version of SC-1. Contrary to what is usually seen in the literature, the presence of yeast extract did not contribute to higher biomass or productivity. Based on these results, we used SC-2 media and target growth rate of $0.13h^{-1}$ for future experiments.

Inoculation cell density and induction time were evaluated as a follow-up study. We targeted 3 inoculation cell densities (1, 3 and 6 OD600nm) and two induction cell densities (35±5 and 18±3 OD600nm). All yeast cultures used the same glucose feed and targeted the same growth rate ($0.13h^{-1}$). As shown in figure 7, all conditions targeting OD600nm of 35 showed higher final cell densities and ethanol production. These trends correlated to ~30% lower volumetric vector yields. Targeting a lower induction cell density helped control ethanol formation and all conditions showed higher vector yields relative to their respective condition induced at an OD600nm of 35. Even though some conditions were induced at half the cell density of their respective control, final cell biomass in galactose was comparable to the conditions induced at higher OD600nm values. We hypothesized that inoculation at higher cell densities would control the formation of byproducts such as ethanol and acetate, and by extension this would impact vector yield. We confirmed this inoculation strategy contributed to lower ethanol formation in culture, and that final vector yields improved 30% in the tested conditions.

We did a final evaluation of the effect of pH, temperature and DO % in vector yield. An I-optimal design which resembled a response surface methodology was used to identify the optimal conditions for our yeast fermentation process. As shown in figure 8, the model generated based on vector yield data suggested that quadratic effects significantly impacted rAAV vector yield in our yeast system, being temperature the factor that exerted the strongest quadratic effect. Temperatures such as 26°C and 34°C impaired per cell vector productivity. An R square value of 0.97 and the results from the analysis of variance confirmed the strength of the model. Based on these results, we were able to identify temperature 30°C, pH 4.8 and DO 10% as the optimal conditions for rAAV-producing, fed-batch fermentation process.

DISCUSSION.

Our initial results in the rAAV2-producing yeast strain confirmed superior performance of the 4-plasmid system compared to the 2-plasmid and the cassette-integrated variants. Our optimal protein expression strategy is based on the use of 2 micron, high-copy number plasmids. This guarantees the presence of 20-50 plasmid copies per cell, compared to gene integration which usually achieves the insertion of 1-3 copies of an expression cassette in the genome (Romanos et al. 1992). We should keep in mind, however, that high plasmid copy number is provided at the expense of low plasmid stability. Although our results showed 30-50% of cells containing the 4 plasmids, we should consider that this number might decrease when attempting high cell density culture strategies. Comparison of performance between 4-plasmid and 2-plasmid system is more complicated to interpret. Although high-copy number plasmids are used in both constructs, Western blot protein expression results showed slightly lower levels of Rep and Cap proteins on the 2-plasmid strains. We speculate that transcription efficiency is impaired in

2-plasmid system constructs, as a result of plasmid saturation with multiple bidirectional GAL1/10 promoters.

The evaluation of codon-optimized genes showed that codon optimization of capsid proteins did not impact vector yield. However, the same study revealed the detrimental effect on rAAV yield by Rep78 codon-optimized gene. These results aligned with previous studies in mammalian cells that suggest Rep78 expression can be associated to cell toxicity and apoptosis (Schmidt et al. 2000).

The media optimization DoE study led to the identification of medium components critical for vector yield. JMP definitive screening design was optimal for this purpose. The fact that the design showed a predicted variance value of 0.04, and *a priori* power values close to 1 for almost all medium components and blocks included, guaranteed the reliability of the study design for this purpose. By including 97 runs, we were able to search for main effects and two-factor interactions between components, preventing confounding effects. The media optimization study revealed a strong positive effect of lysine on rAAV vector yield, whereas zinc sulfate generated a negative impact. For this strain, we hypothesize that SC medium provided minimum levels of lysine for yeast growth, and addition of the amino acid potentially complemented auxotrophic nutrient requirements important for yeast growth and vector yield. Moreover, zinc sulfate is a micronutrient that provides a positive impact for yeast growth and helps alleviate oxidative stress events (Maddox and Hough 1970; Wan et al. 2015). Addition of this zinc salt boosted cell growth, but could have indirectly impacted vector yield due to a growth rate-associated factor.

We succeeded on transferring the fermentation process from 20mL shake flasks to 200mL bioreactors. Our preliminary low cell density, fed-batch fermentation process yielded vector titers comparable to previous shake flasks production runs. Interestingly, a shake flask satellite run of the fermentation run performed poorly. We hypothesized that both

shake flasks and bioreactors suffered the negative effects exerted by byproduct formation, but this effect was partially mitigated by a good control of bioprocessing conditions (pH, temperature, DO %) in bioreactors, which cannot be done in shake flasks. Based on that idea, we believed that a good yeast growth strategy could help control fermentative metabolism and byproduct formation, especially ethanol levels. We then modified the fermentation process to allow for moderate cell growth before galactose induction. Increase of vitamin and amino acids in the SC medium permitted moderate growth. Addition of yeast extract was unnecessary to accomplish that goal. Glucose became the limiting nutrient in SC-2 medium, and was used to control the growth rate in bioreactors. The exponential feeding rate was determined taking several parameters into account such as yeast glucose consumption rate, inoculation cell density, glucose concentration in feed, and target growth rate (Rodríguez-Limas et al. 2011). Feeding rate seemed to play a critical role in the final yields. Adoption of fermentative metabolism as a result of high feeding rates impacted cell growth and vector yield. Feeding rates that target growth rates of $0.13h^{-1}$ or lower promoted respiratory metabolism. Previous reports from other groups attempting high cell density in *Saccharomyces* also selected slow feeding rates for their processes, because the slow, steady yeast growth regulates byproduct formation, and potentially provides more ATP energy as a result of more intermediate molecules going through the TCA cycle and oxidative phosphorylation (Fieschko et al. 1987; Mendoza-Vega et al. 1994a; Mendoza-Vega et al. 1994b).

Modifying inoculation and induction cell density while using the same target growth rate had the objective of controlling the number of duplication cycles, which would likely control byproduct formation and potentially mitigate plasmid loss after subsequent duplication events. This strategy did work at a certain extent; however, it did not translate into significantly higher increments in vector yield. The final response surface study allowed

us to identify critical effects during fed-batch fermentation. Even though the predicted variance profile showed moderate values (0.4-0.6), the response to different parameters was strong enough that results contributed to the generation of a good model. This model identified that quadratic effects for pH, DO % and temperature were significant; the last one being the strongest. As seen during the one-factor-at–a-time experiment in shake flasks, temperature is critical for rAAV vector yield. Other studies have shown that higher temperatures tend to become detrimental for recombinant protein expression, whereas temperatures lower than 28 degrees seem to contribute positively toward expression, folding and stability of recombinant proteins (Mattanovich et al. 2004; Dragosits et al. 2009; Fu et al. 2012).

We developed a novel fed-batch fermentation process for rAAV2 production using *S. cerevisiae*. A basic assessment states all the advantages of using rAAV-producing yeast versus other complex eukaryotic systems (Table 2). The yeast fermentation process uses simple and defined culture medium, which brings big cost savings in production costs. In addition, the use of yeast for rAAV2 production eliminates regulatory hurdles regarding the presence of viral adventitious agents potentially present in mammalian cells. Although yeast generates 5-10 fold less rAAV2 vector than conventional HEK293 mammalian cells, the higher scalability of *S. cerevisiae* in industrial fermenters would allow companies to meet the required vector yields for clinical and commercial purposes. In a hypothetical scenario of a clinical indication that requires doses of $5x10^{13}$ vg/kg, a yeast fermentation process that shows an upstream yield of $1x10^{13}$ vg/L and shows ~10% downstream recovery yield, would require to be scaled to 25,000L to generate enough vector material for ~10 patients. There are big fermentation tanks that can be used for this purpose, therefore this operation should be feasible in industrial settings.

Overall, these studies set the basis of a yeast fermentation process for rAAV production. So far, results of fermentation at low density are comparable to the ones obtained at high cell density. We infer that more studies need to be done to understand how the high cell density approach impacts per cell vector productivity, and what changes would be necessary to optimize cell density and improve rAAV vector yield.

CONCLUSIONS

The aforementioned studies provided clues of optimal plasmid design and fermentation parameters for obtaining high vector yield in our rAAV-producing yeast strain. Based on shake flasks and bioreactor studies, rAAV vector titers ranging around 1-2 x 10^{10} vg/mL were achieved. The yeast exploratory platform is showing encouraging results. However, follow-up studies in fermentation will be important to elucidate drivers of vector yield at high cell density.

Acknowledgments: The authors would like to thank Javier Femenia and Santosh Pande at BioMarin Pharmaceutical Inc for their valuable suggestions throughout the execution of the study.

REFERENCES.

1. Aponte-Ubillus JJ, Barajas D, Peltier J, Bardliving C, Shamlou P, Gold D (2017) Molecular design for recombinant adeno-associated virus (rAAV) vector production. Appl Microbiol Biotechnol. doi: 10.1007/s00253-017-8670-1

2. Backovic A, Cervelli T, Salvetti A, Zentilin L, Giacca M, Galli A (2012) Capsid protein expression and adeno-associated virus like particles assembly in Saccharomyces cerevisiae. Microb Cell Factories 11:124 . doi: 10.1186/1475-2859-11-124

3. Barajas D, Aponte-Ubillus JJ, Akeefe H, Cinek T, Peltier J, Gold D (2017) Generation of infectious recombinant Adeno-associated virus in Saccharomyces cerevisiae. PloS One 12:e0173010 . doi: 10.1371/journal.pone.0173010

4. Daya S, Berns KI (2008) Gene Therapy Using Adeno-Associated Virus Vectors. Clin Microbiol Rev 21:583–593 . doi: 10.1128/CMR.00008-08

5. Dragosits M, Stadlmann J, Albiol J, Baumann K, Maurer M, Gasser B, Sauer M, Altmann F, Ferrer P, Mattanovich D (2009) The effect of temperature on the proteome of recombinant Pichia pastoris. J Proteome Res 8:1380–1392 . doi: 10.1021/pr8007623

6. Fieschko JC, Egan KM, Ritch T, Koski RA, Jones M, Bitter GA (1987) Controlled expression and purification of human immune interferon from high-cell-density fermentations of Saccharomyces cerevisiae. Biotechnol Bioeng 29:1113–1121 . doi: 10.1002/bit.260290911

7. Fu Z, Leighton J, Cheng A, Appelbaum E, Aon JC (2012) Optimization of a Saccharomyces cerevisiae fermentation process for production of a therapeutic recombinant protein using a multivariate Bayesian approach. Biotechnol Prog 28:1095–1105 . doi: 10.1002/btpr.1557

8. Galli A, Della Latta V, Bologna C, Pucciarelli D, Cipriani F, Backovic A, Cervelli T Strategies to optimize capsid protein expression and single-stranded DNA

formation of adeno-associated virus in Saccharomyces cerevisiae. J Appl Microbiol n/a-n/a . doi: 10.1111/jam.13511

9. Maddox IS, Hough JS (1970) EFFECT OF ZINC AND COBALT ON YEAST GROWTH AND FERMENTATION. J Inst Brew 76:262–264 . doi: 10.1002/j.2050-0416.1970.tb03293.x

10. Mattanovich D, Gasser B, Hohenblum H, Sauer M (2004) Stress in recombinant protein producing yeasts. J Biotechnol 113:121–135 . doi: 10.1016/j.jbiotec.2004.04.035

11. Mendoza-Vega O, Hebert C, Brown SW (1994a) Production of recombinant hirudin by high cell density fed-batch cultivations of a Saccharomyces cerevisiae strain: physiological considerations during the bioprocess design. J Biotechnol 32:249–259

12. Mendoza-Vega O, Sabatié J, Brown SW (1994b) Industrial production of heterologous proteins by fed-batch cultures of the yeast Saccharomyces cerevisiae. FEMS Microbiol Rev 15:369–410

13. Rodríguez-Limas WA, Tyo KEJ, Nielsen J, Ramírez OT, Palomares LA (2011) Molecular and process design for rotavirus-like particle production in Saccharomyces cerevisiae. Microb Cell Factories 10:33 . doi: 10.1186/1475-2859-10-33

14. Romanos MA, Scorer CA, Clare JJ (1992) Foreign gene expression in yeast: a review. Yeast Chichester Engl 8:423–488 . doi: 10.1002/yea.320080602

15. Samulski RJ, Muzyczka N (1999) Adeno-associated Viral Vectors. In: The Development of Gene Therapy, Friedmann T.

16. Schmidt M, Afione S, Kotin RM (2000) Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. J Virol 74:9441–9450

17. Tratschin JD, Miller IL, Smith MG, Carter BJ (1985) Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. Mol Cell Biol 5:3251–3260

18. Tratschin JD, West MH, Sandbank T, Carter BJ (1984) A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. Mol Cell Biol 4:2072–2081

19. Urabe M, Ding C, Kotin RM (2002) Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. Hum Gene Ther 13:1935–1943. doi: 10.1089/10430340260355347

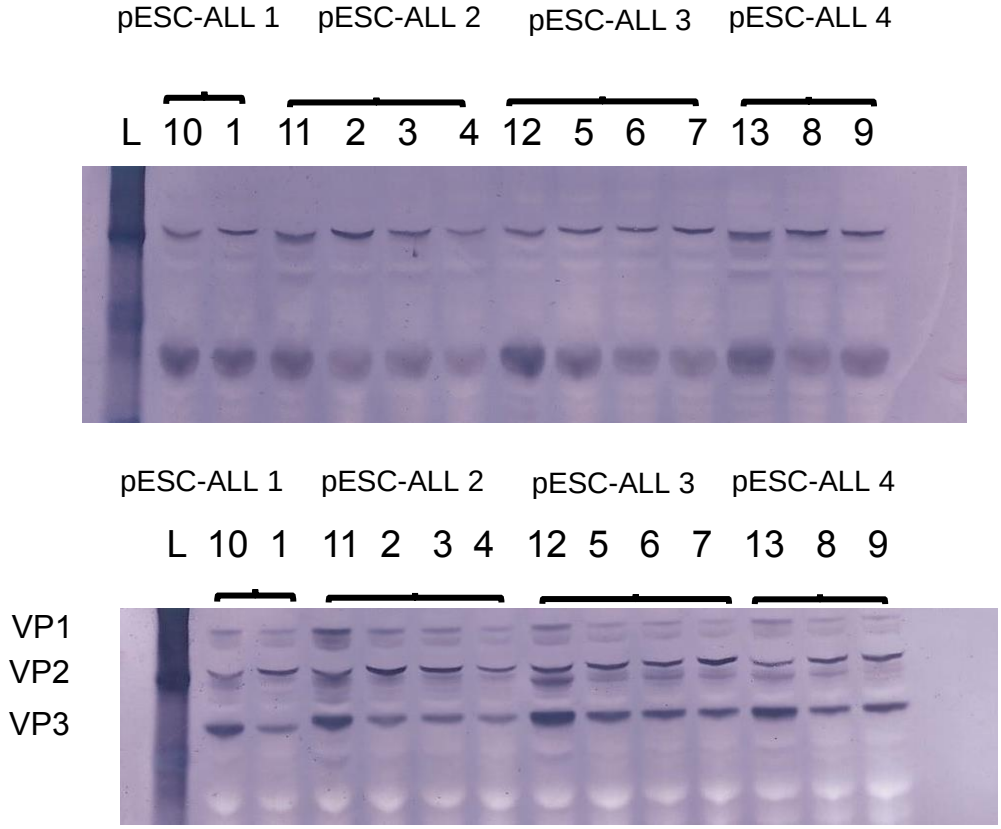
20. Wan C, Zhang M, Fang Q, Xiong L, Zhao X, Hasunuma T, Bai F, Kondo A (2015) The impact of zinc sulfate addition on the dynamic metabolic profiling of Saccharomyces cerevisiae subjected to long term acetic acid stress treatment and identification of key metabolites involved in the antioxidant effect of zinc. Metallomics 7:322–332 . doi: 10.1039/C4MT00275J

FIGURES.

A.

ALL1	ALL2	ALL3	ALL4
VP1	VP1(op)	VP1	VP1
VP2	VP2(op)	VP2(op)	VP2(op)
VP3	VP3(op)	VP3(op)	VP3 (op)
Rep78	Rep78(op)	Rep78	Rep78(op)
Rep52(op)	Rep52(op)	Rep52(op)	Rep52(op)
AAP(op)	AAP(op)	AAP(op)	AAP(op)

B.

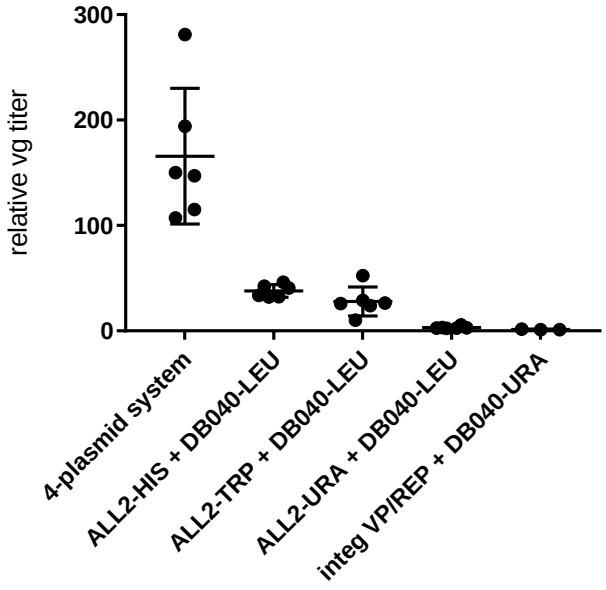


C.

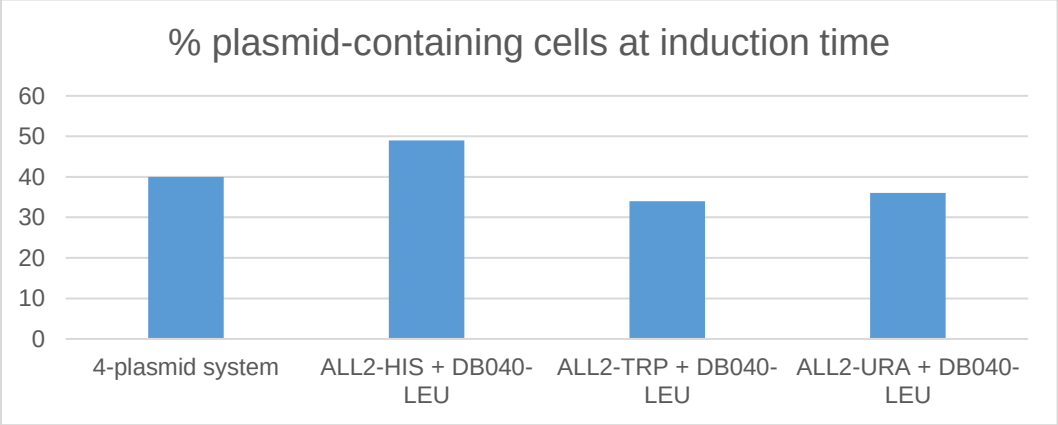
	4-plasmid system	2-plasmid system
	rAAV-GFP/18s rDNA ratio (ddPCR)	
pESC-ALL1	4.1	1.2
pESC-ALL2	3.0	1.2
pESC-ALL3	2.8	1.0
pESC-ALL4	3.9	0.9

Fig 1. rAAV protein expression using 4-plasmid and 2-plasmid system configurations. (A) Description of the 2-plasmid systems assembled. (B) Western blot results of Rep and Cap expression. Numbers 1-9 represent different copies of 2-plasmid versions ALL1-4, whereas numbers 10,11,12,13 represent the equivalent 4-plasmid system version those constructs. (C) ddPCR results of rAAV-GFP/18s DNA ratio analyzed in four 2-plasmid system configurations.

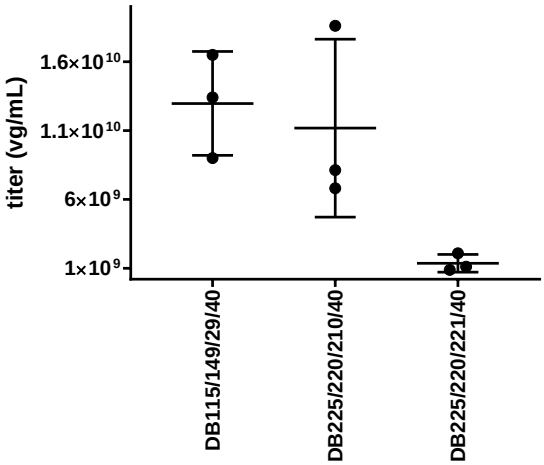
A.



B.



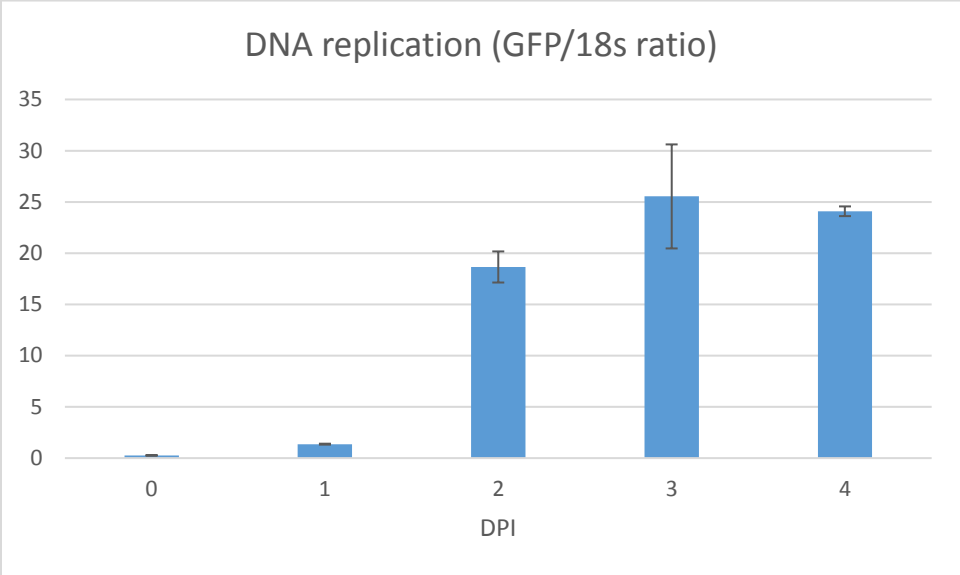
C.



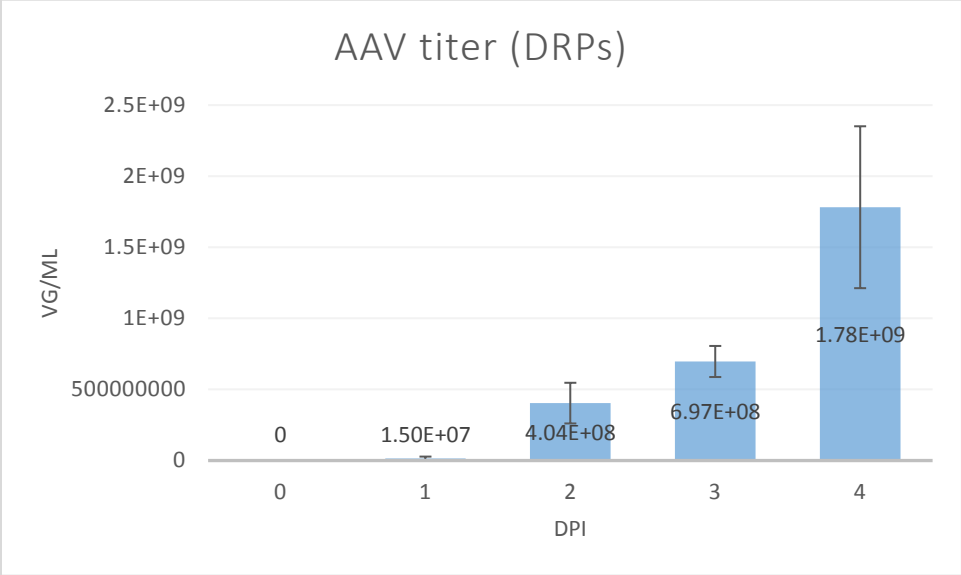
Codon optimization			
VP1	-	+	+
VP2	-	+	+
VP3	-	+	+
AAP	+	+	+
Rep52	+	+	+
Rep78	-	-	+

Fig 2. rAAV vector yield analysis from different 2 and 4-plasmid system configurations. (A) The performance of the standard 4-plasmid system control was compared against 3 2-plasmid system configurations that contained different auxotrophic selection markers, and against a strain that contained all expression cassettes integrated into the genome and only uses one plasmid for the AAV-GFP transgene. (B) Plasmid stability of 4-plasmid and 2-plasmid system configurations at induction time. (C) Evaluation of vector titer using different codon optimized genes.

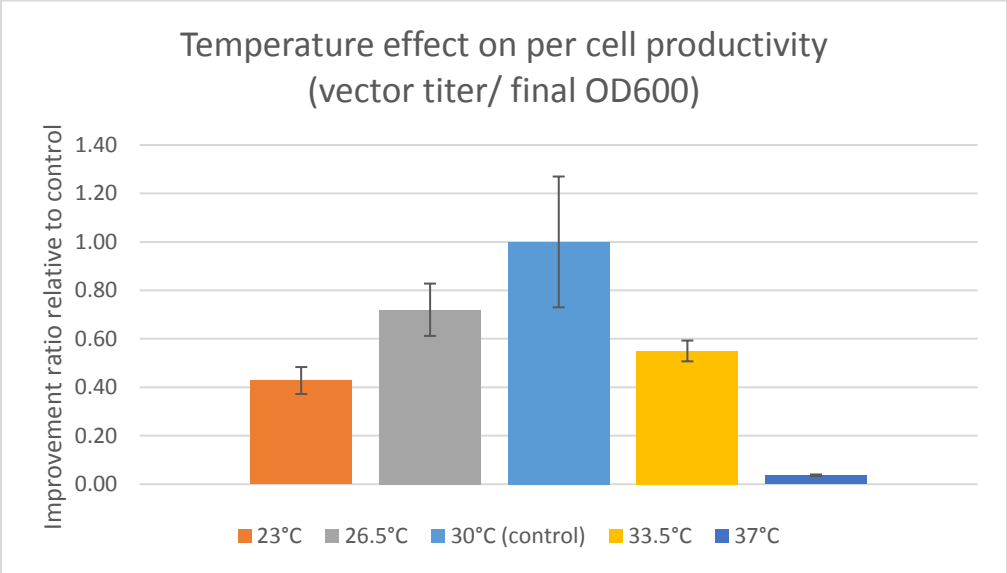
A.



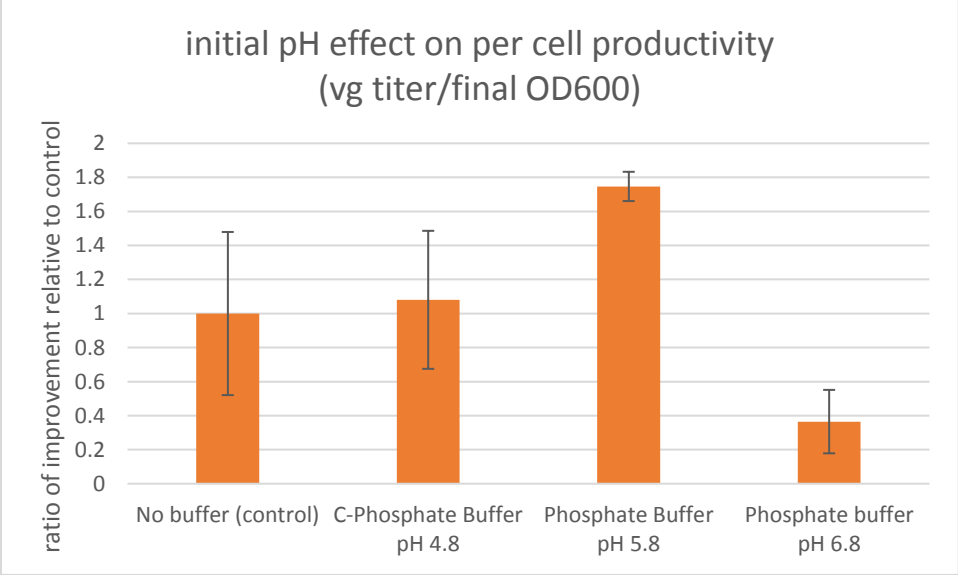
B.



C.



D.



E.

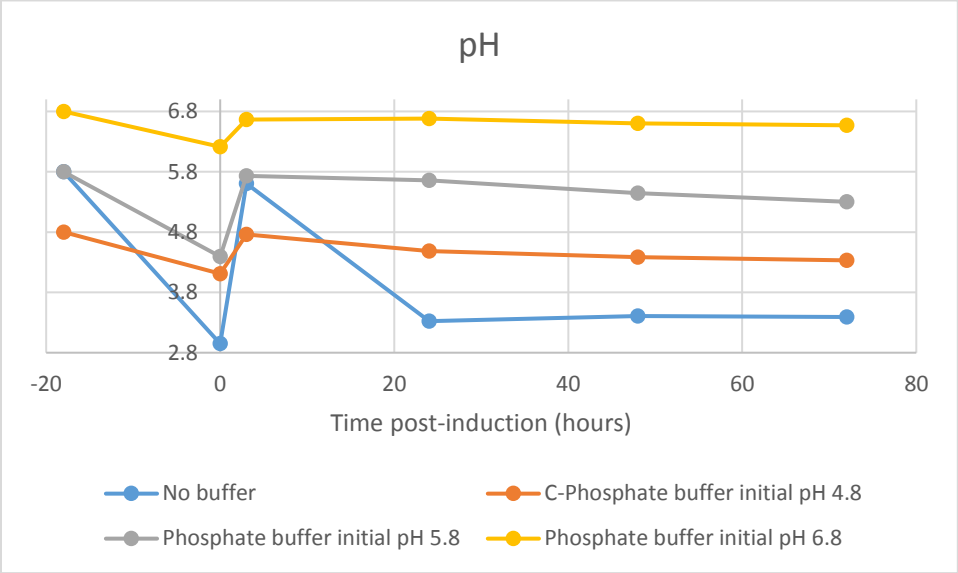
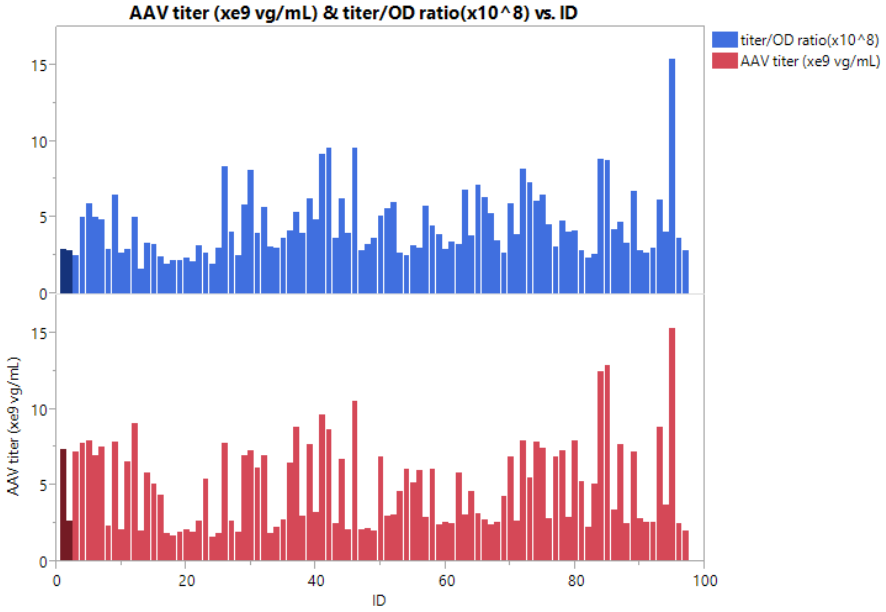


Fig 3. Preliminary production studies performed in shake flasks. (A) Time-course analysis of rAAV DNA replication (rAAV-GFP/18s DNA ratio). (B) Time-course analysis of vector production (Benzonase- resistant rAAV DNA). (C) Evaluation of the effect of temperature on per cell productivity. (D) Evaluation of the effect of pH on per cell productivity. (E) pH trends of evaluated shake flask conditions. Studies were done in triplicates.

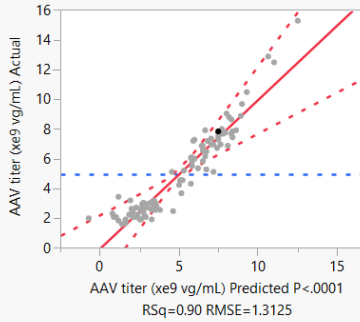
A.



B.

Response AAV titer (xe9 vg/mL)

Actual by Predicted Plot



Summary of Fit

RSquare	0.90176
RSquare Adj	0.790422
Root Mean Square Error	1.312479
Mean of Response	4.968041
Observations (or Sum Wgts)	97

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	51	711.54405	13.9518	8.0993
Error	45	77.51708	1.7226	Prob > F
C. Total	96	789.06113		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	4.9723	0.133283	37.31	<.0001*
Lysine	2.4228723	0.135372	17.90	<.0001*
Zn sulfate	-0.560213	0.135372	-4.14	0.0002*
Block[3]	-0.676883	0.231629	-2.92	0.0054*
Histidine	-0.300745	0.135372	-2.22	0.0314*
Pyridoxin	0.2707447	0.135372	2.00	0.0516
Cysteine	0.2451064	0.135372	1.81	0.0769
Block[1]	-0.4131	0.228509	-1.81	0.0773
Valine	-0.244043	0.135372	-1.80	0.0781
Fe Chloride	0.2410638	0.135372	1.78	0.0817
Choline Cl	-0.240851	0.135372	-1.78	0.0820
myo-Inositol	0.2379787	0.135372	1.76	0.0856

C.

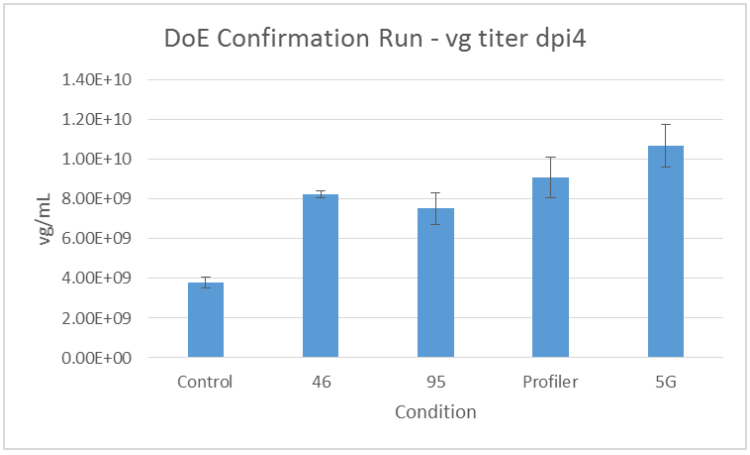
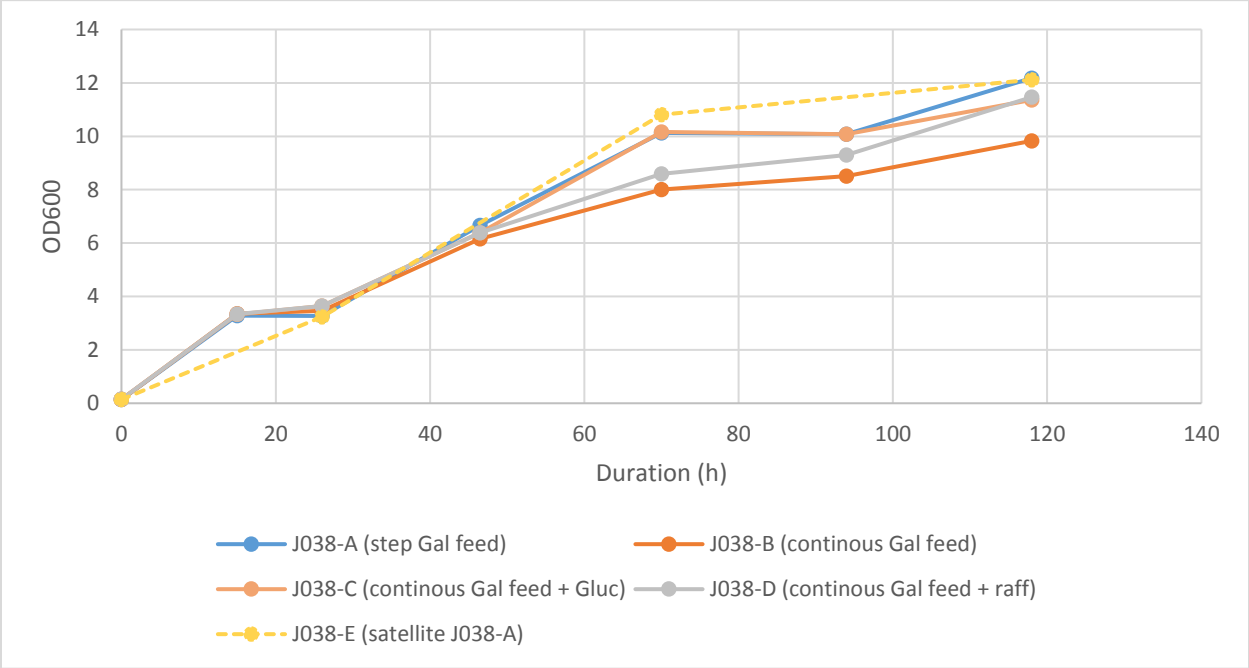


Fig 4. DoE media screening study. (A) Histogram showing all vector yield (vg/mL) and per cell productivity (vg/OD600nm) results obtained along conditions. (B) Predicted model generate by JMP software, which identifies media components that exert a significant effect on vector yield. (C) Confirmation runs of the top conditions, compared to the control condition.

A.



B.

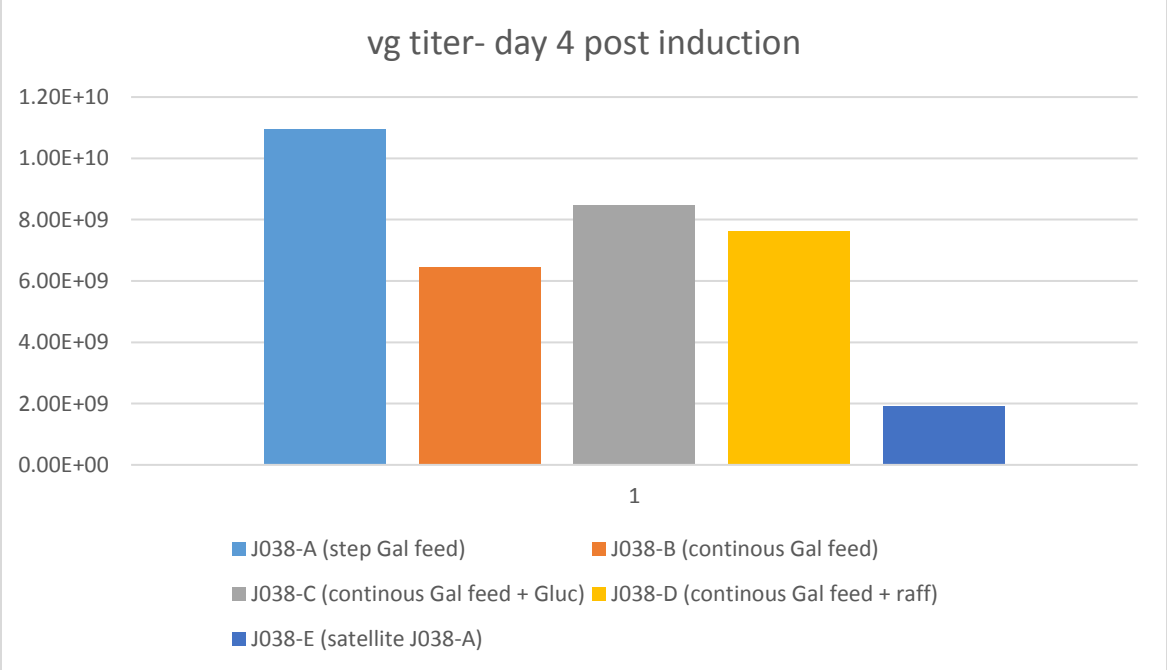
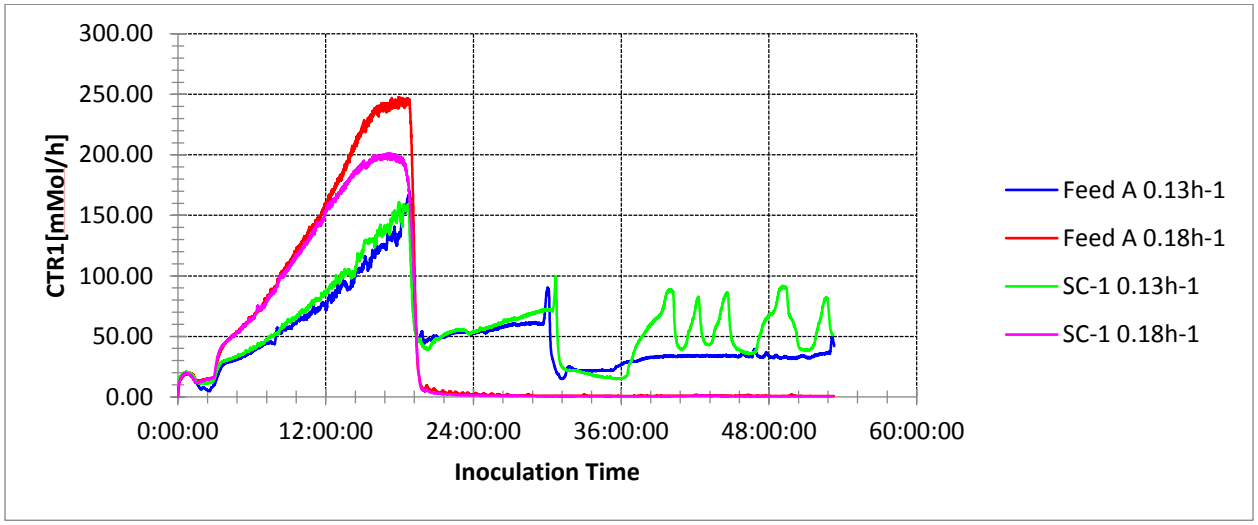
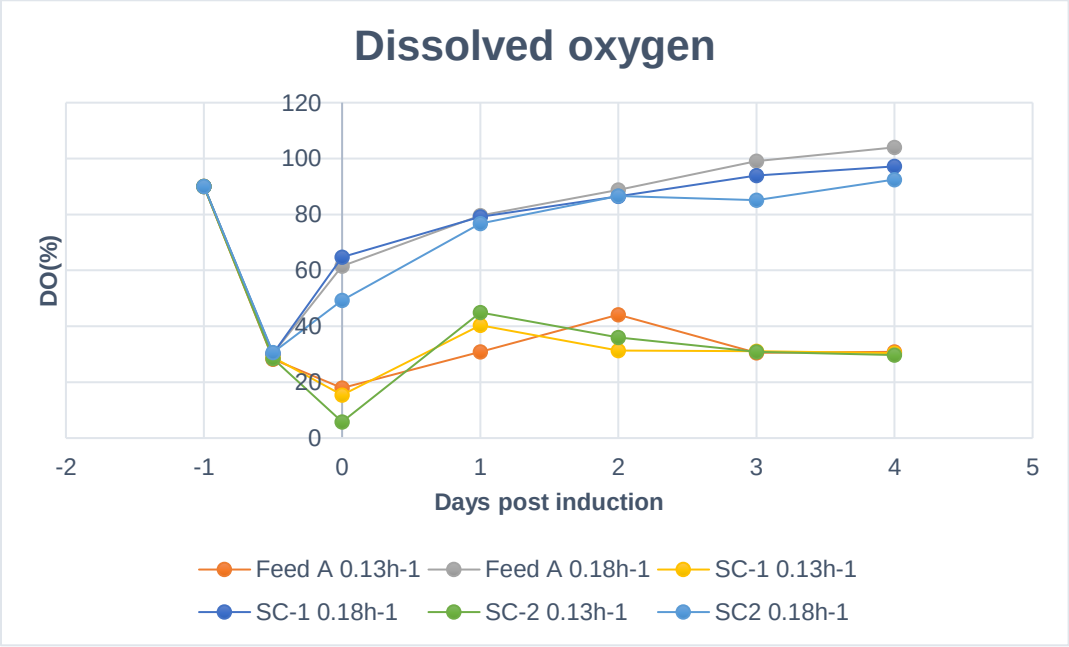


Fig 5. Evaluation of different galactose feed strategies at low cell densities. Five strategies were evaluated in 250ml bioreactors. A satellite run was performed in 125mL shake flask. (A) Cell growth results, and (B), vector yield results at day 4 post induction.

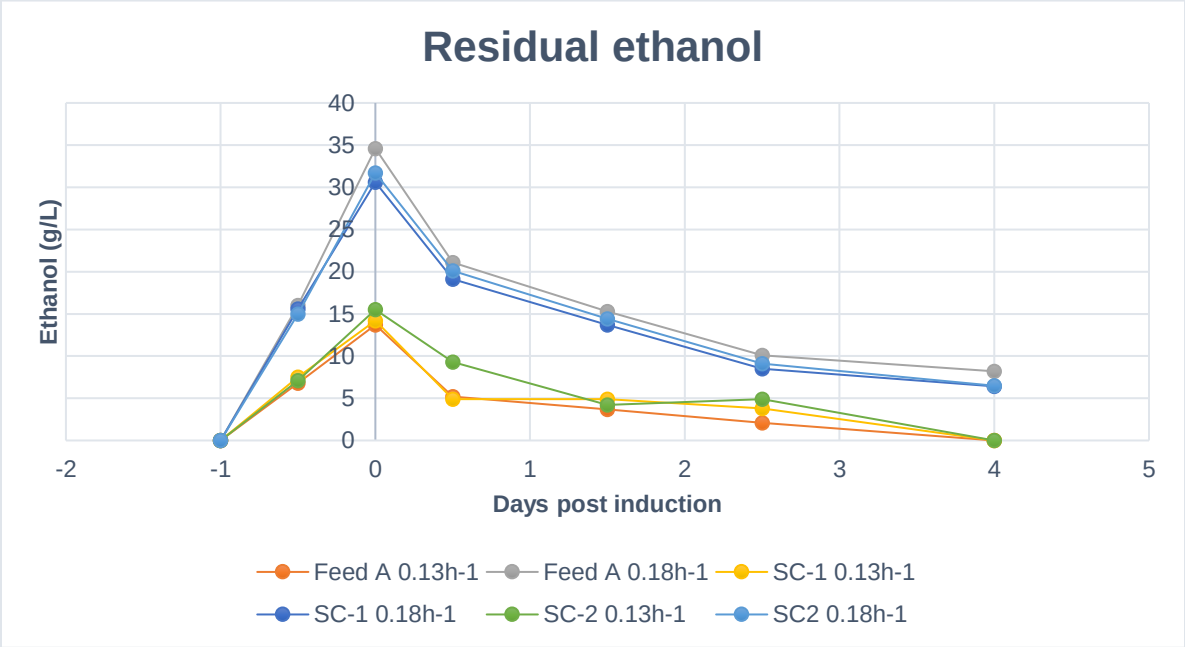
A.



B.



C.



D.

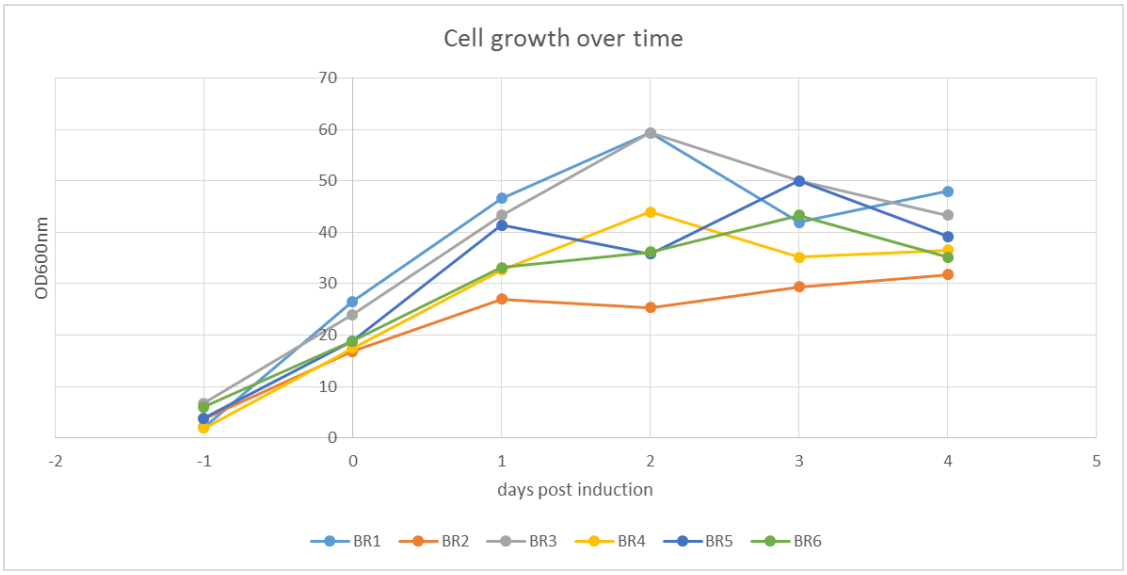
Conditions	Final DCW (g/L)	vector genome titer (vg/mL)
Feed A 0.13h-1	35	5.29E+09
Feed A 0.18h-1	22	<10E+07
SC-1 0.13h-1	40	6.78E+09
SC-1 0.18h-1	21	<10E+07
SC-2 0.13h-1	46	1.13E+10
SC-2 0.18h-1	19	<10E+07

Fig 6. Evaluation of the effect of glucose feed type and glucose feeding rate in yeast cell growth and vector yield. (A) Carbon transfer rate (mMol/h) results from conditions fed with two different media (Feed A vs SC-1) and targeting two different growth rates ($0.13h^{-1}$ vs $0.18h^{-1}$). (B) Cell growth trends of the six tested conditions. (C) Ethanol byproduct formation level for all conditions. (D) Vector yields obtained at day 4 post induction.

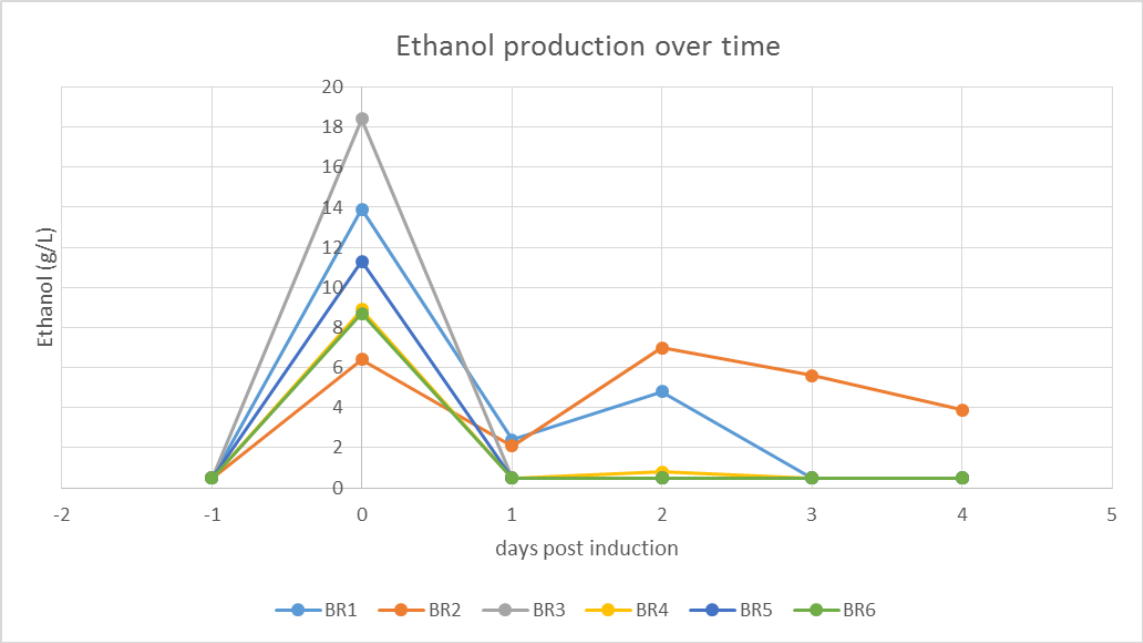
A.

BR	Target growth rate	Inoc cell density	time(h) to induction	time(h) to induction
1	0.13h-1	1.5	24.2	24h
2	0.13h-1	3	18.9	19h
3	0.13h-1	6	13.6	14h
4	0.13h-1	1.5	18.7	19h
5	0.13h-1	3	13.3	14h
6	0.13h-1	6	8.0	9h

B.



C.



D.

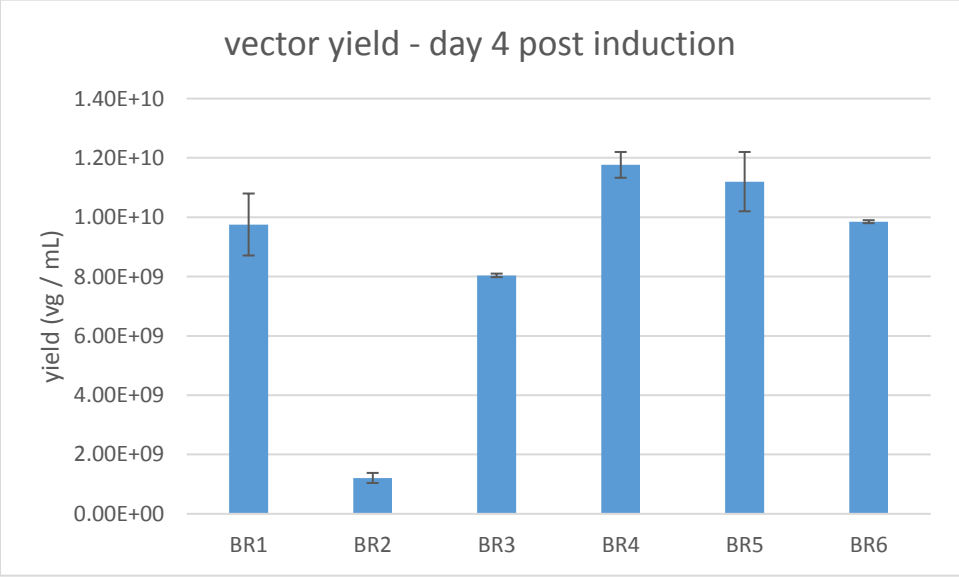


Fig 7. Evaluation of the effect of inoculation cell density and induction cell density on vector yield. (A) table that describes the tested conditions. (B) cell growth trends, (C) ethanol byproduct formation levels. (D) vector yield results at day 4 post induction.

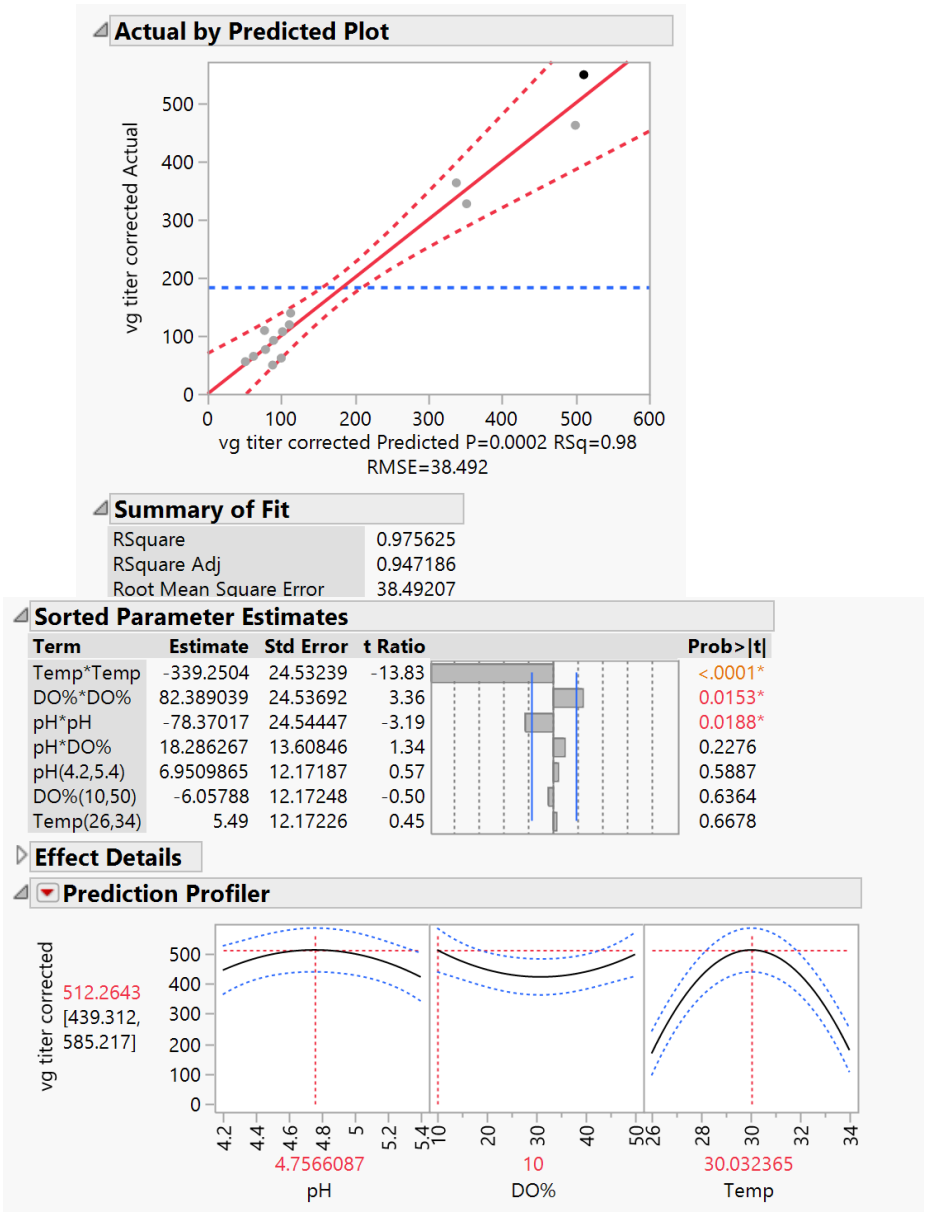


Fig 8. DoE response surface design to evaluate the effect of DO%, temperature and pH on vector yield (vector yield relative values are shown on y-axis).

TABLES

Table 1. List of plasmids used during this study.

Plasmid code	Description
DB155	DB155-pESC(H)_VP3_AAP(opt)_TEF1p-FKBP46-HA
DB149	DB149-pESC(L)_Rep52(2op)_VP1(2)_ADH2p-E2A(op)-HA
DB029	DB029-pESC(T)_Rep78_VP2
DB040	DB040-pAAV-GFP_2mic-URA3
JA001	VP1-VP2-VP3-Rep78-Rep52(op)-AAP(op)
JA002	AAV-GFP
JA001-ALL2	VP1(op)-VP2(op)-VP3(op)-Rep78(op)-Rep52(op)-AAP(op)
JA001-ALL3	VP1-VP2(op)-VP3(op)-Rep78-Rep52(op)-AAP(op)
JA001-ALL4	VP1-VP2(op)-VP3(op)-Rep78(op)-Rep52(op)-AAP(op)
DB225	pESC(H)_VP3(2op)_AAP(2op)_TEF1p-FKBP46
DB220	pESC(L)Rep52op_Gal10E2A_Gal7VP1(2op)
DB210	pESC(T)_Rep78_VP2(2op)
DB221	pESC(T)_Rep78(2op)_VP2(2op)

Table 2. Design matrix of the DoE bioreactor study to optimize pH, temperature and DO%.

Run	pH	DO%	Temp	Blocking
1	4.2	50	34	1
2	4.8	30	34	1
3	5.4	30	30	1
4	4.824	10	30	1
5	4.2	10	26	1
6	5.4	50	26	1
7	4.2	50	26	1
8	5.4	50	34	2
9	5.4	10	34	2
10	5.4	10	26	2
11	4.8	30	26	2
12	4.2	30	30	2
13	4.2	10	34	2
14	4.8	50	30	2

Table 3. Comparison between mammalian HEK293 and potential rAAV-producing yeast expression systems.

HEK293	***S. cerevisiae***
Use of complex, expensive culture medium	Use of simple, affordable culture medium
Cells can be lysed and product harvested from supernatant	Product has to be harvested from the intracellular environment
Challenging scale-up (up to ~2000L)	Easy to scale-up (up to ~100,000L)
Requires production of large amounts of plasmid for N-stage transfection	Transformed clone can be expanded before rAAV production
Regulatory concerns regarding human nature of cell line and presence of adventitious agents	No regulatory concerns regarding human adventitious agents

CONCLUSIONS

1. Overexpression of host cell factors can modulate rAAV DNA replication and vector yield. Our screening study identified several candidates that impacted these two performance outputs significantly. The identified candidates participate in several different biological processes, which suggest that there might be multiple ways these proteins can directly or indirectly influence recombinant vector production. HEM4 and TOP2 proteins were the proteins that showed the highest rate of improvement.
2. Proteomic profiling studies provided more clues about differential biological processes associated to rAAV protein expression. The proteomic profiling results suggest that yeast might undergo protein conformational stress, oxidative stress and environmental stress during rAAV formation. Our protein overexpression strategy validated the functional importance of key host cell proteins during rAAV vector production.
3. Fermentation conditions do impact cell growth and vector productivity. By studying several parameters such as media composition, inoculation cell density, glucose feed, glucose feeding rate, induction length, galactose feeding strategy, pH, temperature and DO%; we were able to identify optimal conditions that promoted vector yields as high as 1.5 $x10^{10}$ vg/mL, which represents a 50 to 100 fold improvement from the results reported in Barajas et al. (2017). More studies need to be done to understand the deficit in per cell productivity when operating at high cell densities.