

RESEARCH ARTICLE

Applied Cellular Physiology and Metabolic Engineering

Transcriptomic features reveal molecular signatures associated with recombinant adeno-associated virus production in HEK293 cells

Yongdan Wang¹ | Qiang Fu² | Yong Suk Lee³ | Sha Sha¹ | Seongkyu Yoon¹ 

¹Department of Chemical Engineering,
University of Massachusetts Lowell, Lowell,
Massachusetts, USA

²Department of Biomedical Engineering and
Biotechnology, University of Massachusetts
Lowell, Lowell, Massachusetts, USA

³Department of Pharmaceutical Sciences,
University of Massachusetts Lowell, Lowell,
Massachusetts, USA

Correspondence

Seongkyu Yoon and Sha Sha, 1 University
Avenue, Lowell, MA 01854, USA.

Email: seongkyu_yoon@uml.edu;
xiami1117@gmail.com

Funding information

National Science Foundation, Grant/Award
Number: 1624684 2100075; National Institute
for Innovation in Manufacturing
Biopharmaceuticals, Grant/Award Number:
70NANB17H002

Abstract

The development of gene therapies based on recombinant adeno-associated viruses (rAAVs) has grown exponentially, so the current rAAV manufacturing platform needs to be more efficient to satisfy rising demands. Viral production exerts great demand on cellular substrates, energy, and machinery; therefore, viral production relies heavily on the physiology of the host cell. Transcriptomics, as a mechanism-driven tool, was applied to identify significantly regulated pathways and to study cellular features of the host cell for supporting rAAV production. This study investigated the transcriptomic features of two cell lines cultured in their respective media by comparing viral-producing cultures with non-producing cultures over time in parental human embryonic kidney cells (HEK293). The results demonstrate that the innate immune response signaling pathways of host cells (e.g., RIG-I-like receptor signaling pathway, Toll-like receptor signaling pathway, cytosolic DNA sensing pathway, JAK-STAT signaling pathway) were significantly enriched and upregulated. This was accompanied by the host cellular stress responses, including endoplasmic reticulum stress, autophagy, and apoptosis in viral production. In contrast, fatty acid metabolism and neutral amino acid transport were downregulated in the late phase of viral production. Our transcriptomics analysis reveals the cell-line independent signatures for rAAV production and serves as a significant reference for further studies targeting the productivity improvement in the future.

Abbreviations: AAV, adeno-associated virus; ALDH6A1, aldehyde dehydrogenase 6 family member A1; ARSB, arylsulfatase B; BCAA, branched-chain amino acid; CCL5, CC chemokine ligand 5; cGAS, cyclic GMP-AMP synthase; CHO, Chinese hamster ovary; CLN6, ceroid-lipofuscinosis neuronal protein 6; CXCL9, CXC chemokine ligand 9; DAL, DNA-dependent activator of IFN-regulatory factors; DEG, differentially expressed gene; DOE, design of experiment; EIF2 α , eukaryotic translation initiation factor 2A; ER, endoplasmic reticulum; GADD34, growth arrest and DNA damage inducible 34; GO, gene ontology; GOI, gene of interest; HEK, human embryonic kidney; HSPA6, heat shock protein family A (hsp70) member 6; HSV, Herpes simplex virus; ICPO, infected cell protein no. 0; IFIT, interferon-induced protein with tetratricopeptide repeats; IFITM, interferon-induced transmembrane; IFN, interferon; IFNAR, interferon alpha and beta receptor subunit 1; IFNB1, interferon beta 1; IKK2, inhibitor of kappa-B kinase; IRF, interferon regulatory factor; ISG, interferon-stimulated genes; JAK, Janus kinase; MDA5, melanoma differentiation-associated protein 5; NLRCS5, Nod-like receptor family card domain containing 5; OAS, oligoadenylate synthetase; PEI, polyethylenimine; PERK, protein kinase R-like endoplasmic reticulum kinase; PLS-DA, partial least squares-discriminant analysis; POLR3C, RNA polymerase III subunit C; RIG, retinoic acid-inducible gene; RIN, RNA integrity number; RSAD2, radical S-adenosyl methionine domain containing 2; RSEM, RNA-Seq by expectation maximization; SLC43A2, solute carrier family 43 member 2; STAT, signal transducer and activator of transcription; TBK1, tank-binding kinase 1; TLR, toll-like receptor; TMEM106B, transmembrane protein 106B; TRIF, TIR domain containing adapter-inducing interferon- β ; TRIM25, tripartite motif containing 25; UPR, unfolded protein response; XBP1, X-box binding protein 1.

Yongdan Wang and Qiang Fu equally contributed to this article.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Biotechnology Progress* published by Wiley Periodicals LLC on behalf of American Institute of Chemical Engineers.

KEYWORDS

HEK293 cells, pathway analysis, rAAV production, transcriptomics, transient transfection

1 | INTRODUCTION

As more gene therapy drugs enter the market, AAV-mediated gene therapy has attracted particular interest due to its unique therapeutic advantages, including non-pathogenicity in humans, the ability to target different tissues with various serotypes, and long-term efficacy.¹⁻³ The most common method for clinical use and commercial production is the triple plasmid-based transient transfection of HEK293 cells in order to produce rAAV.⁴ During rAAV production, three plasmids (GOI plasmid, RepCap plasmid, and helper plasmid) are co-transfected into HEK293 cells. This transient transfection process enables rAAV production for a limited time without integrating viral DNA into the host cell genome.⁵⁻⁷ The major challenge of the rAAV manufacturing process is its limited production capability, which cannot meet the demand.⁸ This increases the cost of AAV-mediated gene therapies.⁸

Several strategies have been attempted to improve rAAV productivity. Transient transfection is a complex process, so parameter optimization is one of the most commonly applied approaches.⁹⁻¹¹ The design of experiment (DoE) was conducted to improve the process.¹¹ However, these methods may be limited to specific cell lines or media. One mechanistic model was reported for the rAAV viral vector biosynthesis,¹² but an overall insufficient mechanism-directed understanding of transfection-based rAAV production still limited our ability to capture cellular or molecular features for more efficient production. For example, the rAAV production process is highly reliant on host cells: It requires the host's cellular machinery for its own genome replication, and this genome replication is regulated by the host cell's DNA replication system. However, very limited knowledge of AAV genome replication has been gained so far, especially from the perspective of the host cell.

The exploration of transcriptomes is an important approach to characterizing features of the host cell in terms of viral production. Previous studies have applied microarray and mRNA sequencing (RNA-seq) analysis for specific gene expression and cellular pathway identification in CHO cells^{13,14} and in HEK293 cells,¹⁵ thus consolidating our knowledge on pathway regulation in response to the production of therapeutic drugs. RNA-seq has become a more popular approach to exploring the HEK293 transcriptome, due to the limitations of microarray analysis and the availability of human and HEK293 cell transcriptome databases or reference materials.^{16,17} In a recent study, RNA-seq was used to investigate various engineered viral production¹⁸⁻²⁰ and virus infection mechanisms.²¹⁻²³ However, the use of RNA-seq analysis of the transcriptome to characterize HEK293 cells for rAAV production has not been reported.

In this study, we compare the transcriptome of an AAV-producing group (referred to as the *viral-producing* group) to a non-producing group (the *negative control* group) from two parental HEK293 cell lines on different post-transfection days. These HEK293 cell lines

originated from different sources and were adapted and cultured in their respective media. The differentially expressed genes were analyzed after obtaining the global transcriptome. First, the genes were screened by *fold change* (FC) and by *p*-adj. value thresholds. Then, their biological functions were further explored to understand their relevance to rAAV production. This work presents an attempt to understand the changes in host cell physiology for rAAV production. Based on the results of this study, potential medium supplementation strategies and cell line engineering strategies to enhance viral production are also proposed for future investigations.

2 | MATERIALS AND METHODS

2.1 | Plasmid preparation and cell culture

Escherichia coli stabs contained one backbone plasmid pcDNA3.1/Zeo (Plasmid #V86020, Invitrogen, USA) in addition to three AAV-related plasmids: pAdDeltaF6 (Plasmid #112867, Addgene, USA), pAAV2/2 (Plasmid #104963, Addgene, USA), and AAV-CMV-GFP (Plasmid #67634, Addgene, USA). The bacteria were amplified in Luria-Bertani (LB) broth media with 100 µg/mL ampicillin (Sigma-Aldrich, USA). The plasmids were extracted and purified using the Zymo Maxiprep kit (Zymo, USA). The purified plasmids were then sterilized using a 0.22 µm PES sterile syringe filter (VWR, USA). The quality of plasmids was inspected based on A230/260, A260/280 using Nanodrop (Thermo Fisher, USA), and DNA electrophoresis.

This study used two different media (AMBIC 1.0 HEK293 in-house media and BalanCD HEK293 media [Irvine, USA]) and two different sources of HEK 293 cells: (a) HEK293 cells from ATCC ([HEK-293] CRL-1573, ATCC) adapted in AMBIC media supplemented with 4 mM GlutaMax™ (Thermo Fisher, USA) and (b) HEK293 cells from Mass Biologics adapted in BalanCD HEK293 media with 4 mM glutamine (Thermo Fisher, USA) added. Both sources were cultured in shake flasks (5% CO₂, 37°C, and 125 rpm). For rAAV production, the transfection process utilized 30 mL of cell culture in a 125 mL shake flask.

2.2 | Transient transfection-rAAV vector production

Cells in the logarithmic growth phase were resuspended with fresh media to target predetermined cell density on the day of transfection. These cells were then transfected with three plasmids (pHelper, pAAV2/2, and pGol) and the transfection reagent PEIpro® (PolyScience, USA). Transfection was conducted as follows: Three sterilized plasmids were mixed carefully in the fresh media with 10% of the cell culture volume. The PEIpro was then added to the plasmid

diluted media. After 5–30 min of incubation, the PEIpro–DNA complex medium was added to the cell culture shake flasks. The culture was conducted in a shaker incubator (5% CO₂, 37°C, and 125 rpm). The specific transfection conditions are shown in Tables S1 and S2. To achieve statistical significance for RNA-seq data, biological triplicates were used for each cell culture condition.

2.3 | Cell count, viability, and AAV genome titer measurement

Cell count and viability were evaluated daily using a Cedex HiRes Analyzer (Roche Life Science, USA). The crude harvest was lysed using three freeze–thaw cycles (i.e., frozen in a dry ice–ethanol bath for 2 min and then thawed in 37°C water for 3 min) with vortexing at each thaw. Treatment with MgCl₂ (Sigma-Aldrich, USA) and Benzonase® (New England Biolabs, USA) was then performed, followed by 1 h of incubation and 40 min of centrifugation at 4100g at 4°C. The collected supernatant was treated with DNase I and DNase reaction buffer to eliminate any host proteins or unpackaged GFP (New England Biolabs, USA). This was followed by 1 h of incubation at 37°C. Then, the sample was treated with proteinase K to open the viral capsids (New England Biolabs, USA), followed by 1 h of incubation at 55°C, followed by 10-min inactivation process at 95°C. The sample underwent 125× serial dilution for the quantitative polymerase chain reaction (qPCR) reaction using CFX Real-Time PCR Detection system (Bio-Rad, USA).

A PrimeTime® qPCR Probe assay (Integrated DNA Technologies, USA) was designed using the PrimerQuest® Design Tool (Integrated DNA Technologies, USA). Each 20 µL sample of the qPCR reaction mixture contained 10.0 µL of TaqMan™ Universal PCR Master Mix, 1.0 µL of primer–probe mix (500 nM primer and 250 nM probe), 6.0 µL of molecular biology-grade water, and 3.0 µL of the diluted sample. The cycling conditions were 20 s at 95°C followed by 39 cycles of two-step thermal cycling (3 s at 95°C followed by 30 s at 60°C). For each run, the negative control (i.e., no primer–probe mixture), the no-template control (NTC), and the positive control (i.e., linearized GFP plasmids) were used along with six rAAV5 reference standards in triplicates (MassBiologics, USA) to achieve semi-quantification of viral titers. The primer and probe sequences used for the genome titer were:

Forward: 5'- GAA CCG CAT CGA GCT GAA -3'.

Reverse: 5'- TGC TTG TCG GCC ATG ATA TAG -3'.

Probe: /56-FAM/ATC GAC TTC/ZEN/AAG GAG GAC GGC AAC/3IABkFQ/.

2.4 | RNA isolation and sequencing

Approximately 1 × 10⁶ cells were centrifuged from cell suspension at 1000 rpm for 5 min. The cell pellets were washed with iced PBS twice and then stored at –80°C. Total RNAs were extracted using RNeasy

Mini Kit (QIAGEN, USA) with DNase digestion (QIAGEN, USA) following the manufacturer's instructions. RNA concentrations were quantified using a Qubit fluorometer (ThermoFisher Scientific, USA). The integrity of the total RNA was measured using a Bioanalyzer (Agilent, USA). RNA samples with RIN value greater than 7 were qualified for later cDNA library preparation. Messenger RNA purification and library preparation were performed under the guidance of Universal Plus™ mRNA-Seq with NuQuant® (Tecan Group Ltd., Switzerland). The adaptor-ligated libraries were quantified by quantitative PCR using a CFX Real-Time PCR Detection System (Bio-Rad, USA) and sequenced on a NextSeq 500 System (Illumina, USA) using a NextSeq 500 High-Output v2 Kit (at 150 cycles). Raw sequencing data (in.bcl format) were de-multiplexed according to their barcodes, then converted to FASTQ files using bcl2fastq2 Conversion Software (Illumina, USA). Approximately 10 million 150-base pair (bp) reads (pair-end) were generated for each sample (see Tables S4 and S5). The library size was normalized by DESeq2 in the later bioinformatics pipeline.

2.5 | Alignment to plasmid sequences and the human genome

AAV reference sequences (including *Rep* and *Cap* genes) are detailed in the Supporting Information. The annotation GTF file *Human genome* was used as the global reference. FastQC (Babraham Institute, UK) was used to evaluate the quality of the sequence reads. Artificial adaptors were removed, and reads with a quality below 15 per bp sequence were cropped using Trimmomatic.²⁴ Fewer than 2% of the total reads were aligned to the common RNAs, and these were excluded. Reads from each sample were aligned to PAdeltaF6, *rep* and *cap* plasmid sequences using Bowtie2 as a aligner.²⁵ The remaining reads were aligned to the *Human genome* reference transcriptome using RNA-Seq by expectation maximization (RSEM)²⁶ to quantify transcript abundance using RNA-Seq data. Around 80% of the reads were aligned from each sample to the human genome. All bioinformatics data uploading was performed using Linux command lines in the Massachusetts Green High Performance Computer Center (MGHPCC).

2.6 | Differential gene expression, function analysis, pathway enrichment, and statistical analysis

First, principal component analysis (PCA) and *partial least-squares discriminant analysis* (PLS-DA) in SIMCA (Sartorius, USA) were used to cluster samples with the normalized expected counts output from RSEM. Then, those expected counts were imported to the DEBrowser, an interactive interface for data examination and differential expression.²⁷ DESeq2²⁸ was used to perform differential expression (DE) analyses between different sample groups. A list of differentially expressed genes was generated and filtered according to the fold change between the viral-producing groups and the negative control groups and *p*-adj. values.²⁹ The web-based tool Metascape and

DAVID were used for gene ontology (GO) enrichment and KEGG pathway analysis.³⁰ Origin software (OriginLab, USA) was used for plot generation.

2.7 | Quantitative polymerase chain reaction for relative gene expression quantification

Synthesized cDNA library samples (the same sample used for the RNA sequencing), were diluted 10 times, amplified, and detected in a qPCR using universal SYBR Green (Thermo Fisher, USA) protocol. The primer sequences were shown in the Table S10. The comparative cycle threshold ($2^{-\Delta\Delta C_t}$) method was used to analyze the relative gene expression level changes. The transcript level of each gene was

normalized to the housekeeping gene *GAPDH*. The relative gene expression for each sample was quantified in technical duplicates.

3 | RESULTS AND DISCUSSION

3.1 | Post transfection cell growth, viability, and viral titer

For the viral-producing condition, parental HEK293 cells were transfected with helper, RepCap, and GOI plasmids. Two production systems were studied using the BalanCD and AMBIC media and their specific cell lines. Both cell culture systems lasted for 72 h after transfection. For the non-producing condition, parental HEK293 cells were

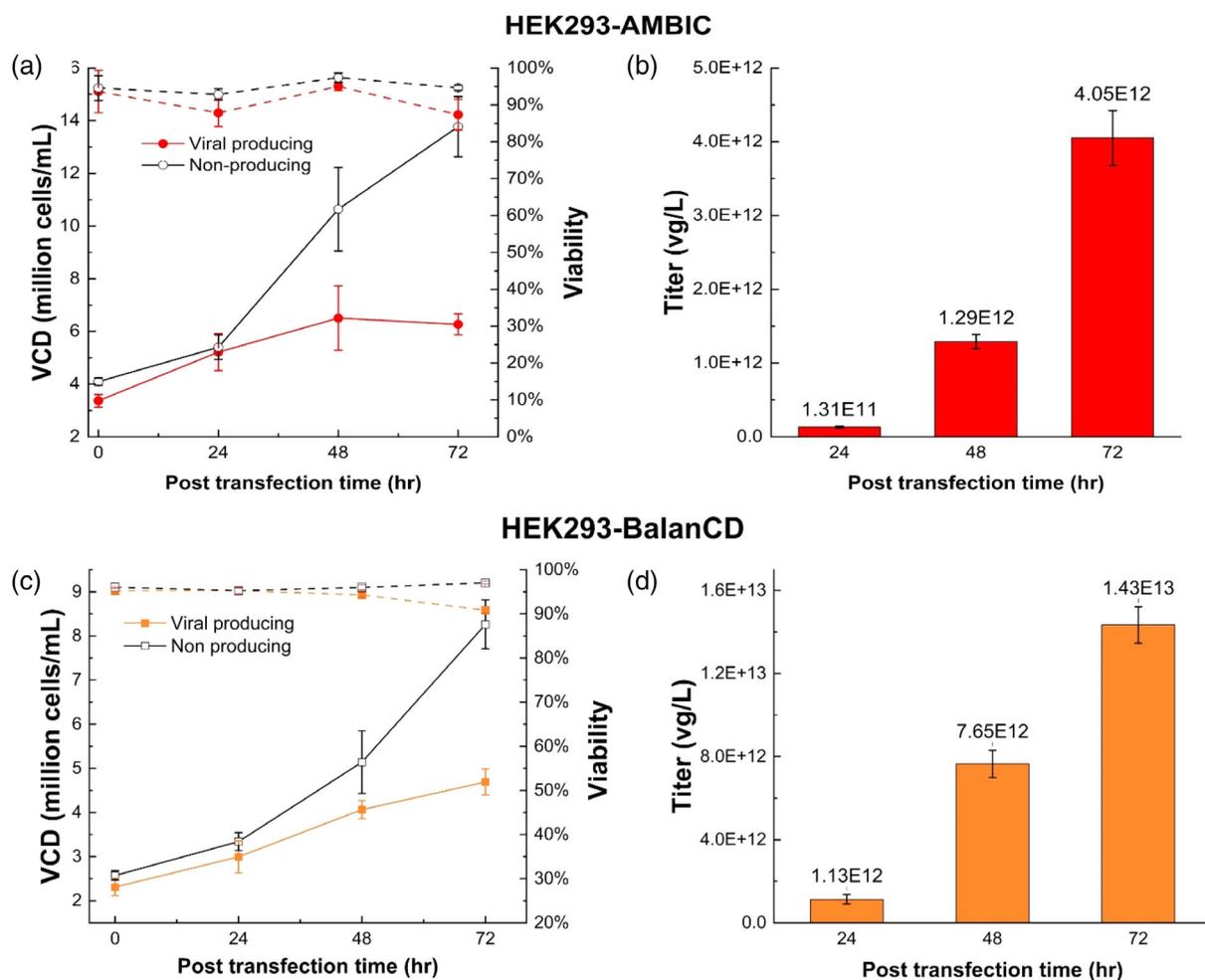


FIGURE 1 Cell growth profile and genome titer for both AMBIC and BalanCD cell cultures post-transfection. Left y axis in (a) and (c) represents viable cell density, expressed in million cells/mL. The right y axis represents viability expressed as a percentage. The y axis in (b) and (d) represents the viral titer expressed in viral genome per liter (vg/L). The closed circle and square represent the viral-producing condition. The open circle and square represent the non-producing condition. Data shown are the average of the triplicate cell cultures. The error bar represents the standard deviation among biological triplicates

transfected with backbone-sequence plasmids at a total mass of DNA equal to the amount used in the viral-producing condition. Each condition was conducted in biological triplicates. Figure 1 shows the cell growth, viability, and viral titer for both production systems. With a similar cell density prior to transfection, post-transfection cell growth and viability in the viral-producing condition were lower than in the non-producing condition for both systems. This phenomenon might be explained by the shift of nutrients and energy from supporting cell growth to rAAV production. At 72 h, the production conditions using the BalanCD and AMBIC systems reached a genome yield of 1.43×10^{13} vg/L and 4.05×10^{12} vg/L, respectively. The cell line adapted to the BalanCD medium achieved a virus titer three times higher than the cell line in the AMBIC medium. Specific productivities (vg/cell) for both viral producing conditions were calculated and shown in Table S3. As expected from the total genome titer, Q_p of HEK293-BalanCD cell culture was continuously higher than that of HEK293-AMBIC cell cultures over time, achieving peak specific productivity 48 h post transfection. Specific differences among cell lines themselves and different components supplemented in media both contributed to titer variations between the two production systems. During this experiment, cell pellet samples were taken 24 h (“D1”),

48 h (“D2”), and 72 h (“D3”) after transfection. These samples were used in the transcriptomic study discussed below.

3.2 | Transcriptome overview differentiating the viral-producing and non-producing conditions

Figure 2 shows a cluster analysis using both PCA and PLS-DA for daily gene expression, which is the output of the normalized transcript counts from RSEM, for the viral-producing and non-producing conditions of each production system. PLS-DA, a supervised statistical method, was utilized and mainly analyzed here as it provided more efficient and better separation with defined and known groups of samples compared to unsupervised PCA analysis. PLS-DA has been recommended for use in omics data analysis since it can successfully reduce data dimensions and also be aware of the classification.³¹ In our case, each biological triplicate set was defined as one distinct class in PLS-DA analysis. RNA-seq data was missing for one of the non-producing group samples in the BalanCD (NC2D2) system, so it was excluded from PLS-DA analysis. The first two main principals, shown as R2x [1] and R2x [2] in each plot in Figure 2c,d, accounted for about

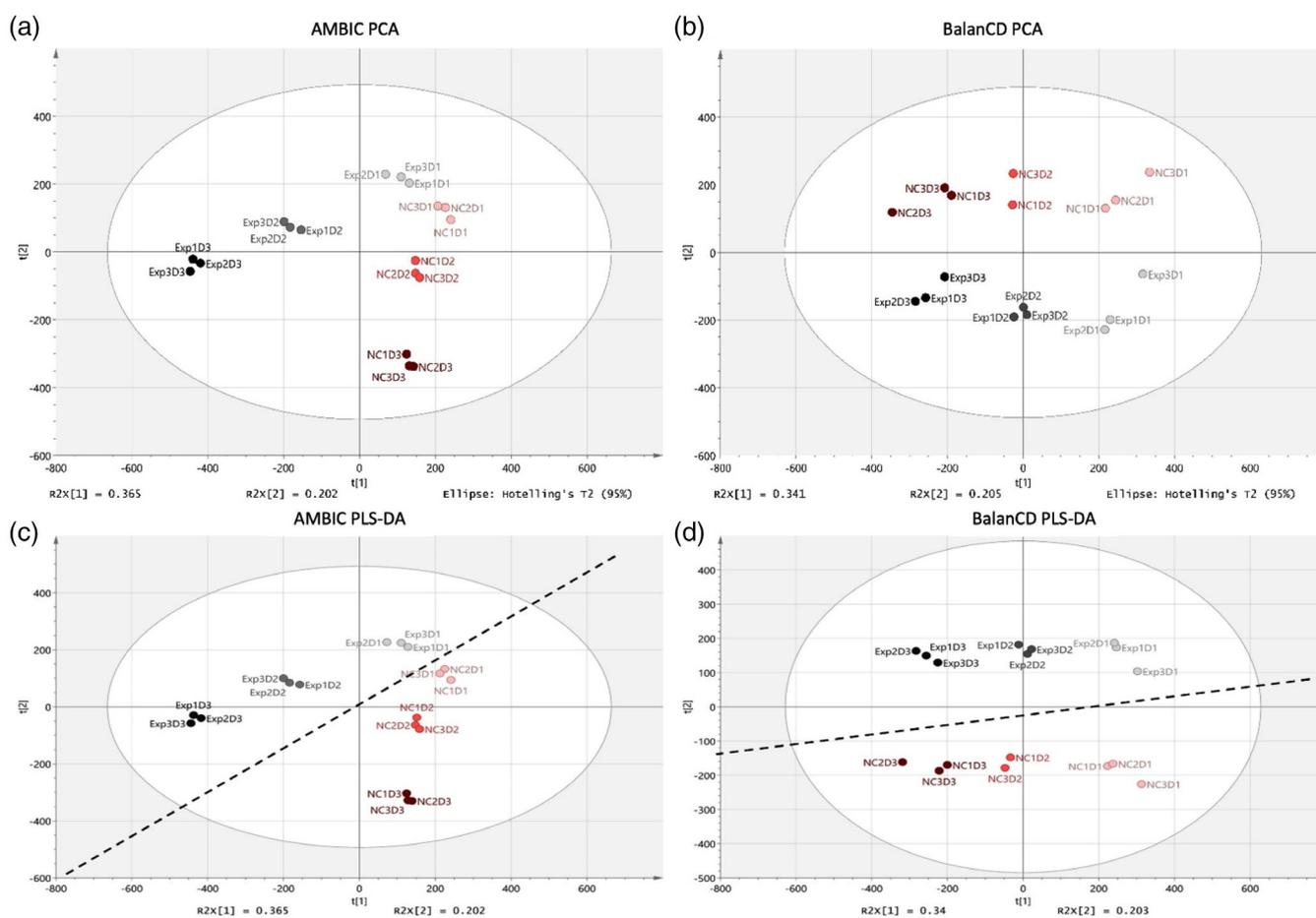


FIGURE 2 Principal component analysis and PLS-DA plots for both AMBIC and BalanCD cell cultures on different post-transfection days. Diagonal lines separate viral-producing (Exp) and non-producing (NC) groups for PLS-DA plots. Post-transfection times are represented as D1 (24 h), D2 (48 h), and D3 (72 h)

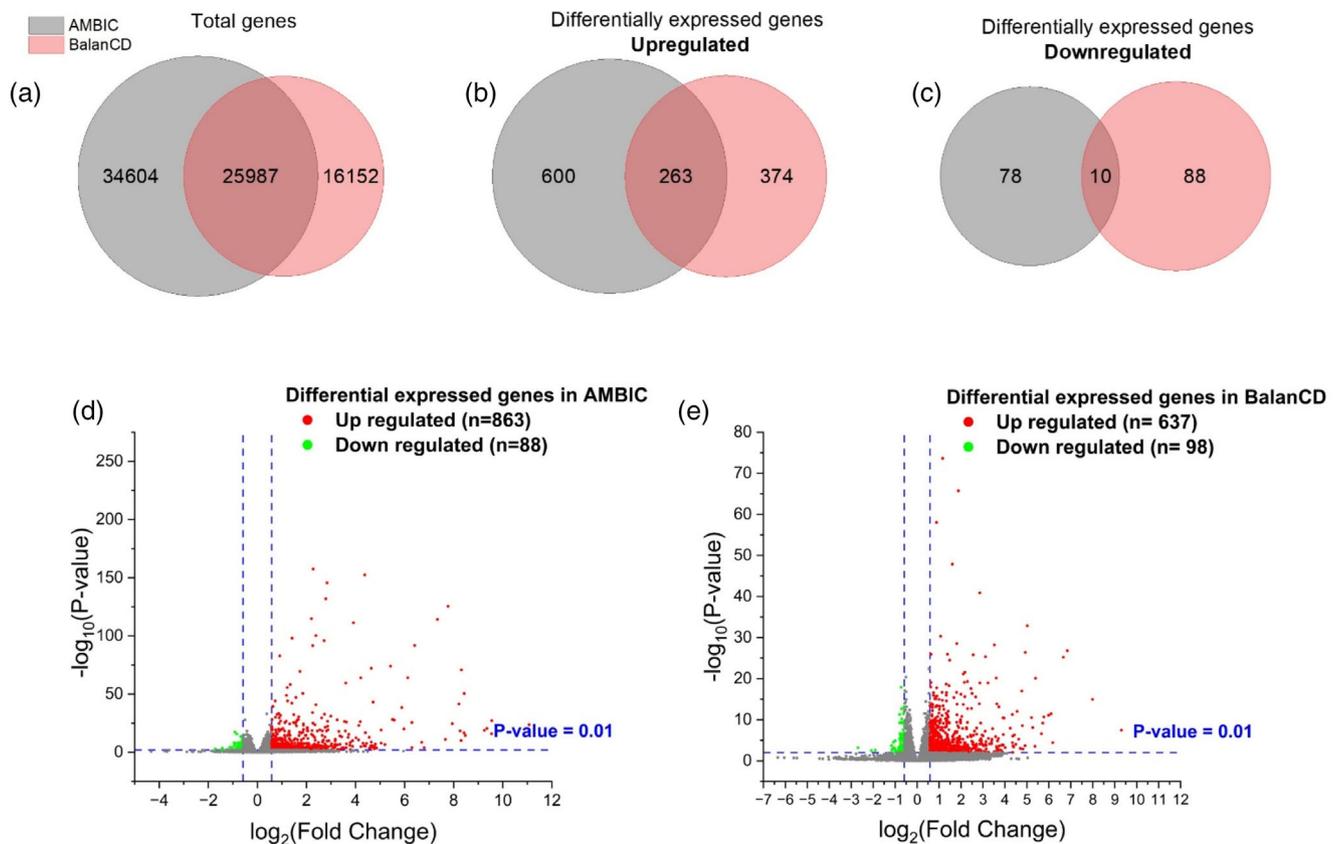


FIGURE 3 Venn diagram of the overall gene expression (a) and the differentially expressed genes (b) and (c). The overlapping region shows the number of genes shared by the AMBIC and BalanCD cell cultures. Volcano plots for differentially expressed genes in the AMBIC (d) and BalanCD cell cultures (e). The fold change threshold was set to 1.5, and the p -adj. value threshold was set to 0.01. The upregulated genes are represented by red dots, the downregulated genes are represented by green dots, and nonsignificant genes are represented by gray dots

55% of the variability in the data. It was noted that the biological replicates for each condition were clustered, and the viral-producing and non-producing groups could be separated by diagonal lines. This result indicates a distinct separation between the viral-producing and non-producing conditions, and the distance between the two conditions increased as the transfection continued over time from the beginning to harvest. This indicates a growing difference between the two groups in terms of cellular physiology after transfection. Further analysis was conducted to screen the differentiated gene expressions and to identify the host cell physiological changes accompanied with rAAV production.

The cellular mRNA from D1, D2, and D3 post-transfection were sequenced for both systems. A total of 60,591 and 42,141 genes were identified in the AMBIC and BalanCD cell cultures respectively. It was found that 25,987 genes overlapped between the two cell lines (see Figure 3a). Next, the differentially expressed genes (DEGs) were identified using DESeq2 analysis between the viral producers and non-producers, including all post-transfection time points, with a threshold of 1.5-fold changes and a p -adj. value < 0.01 . As shown in Figure 3b,c, the viral-producing condition using the AMBIC medium

resulted in 863 upregulated genes and 88 downregulated genes. The viral production system using the BalanCD medium resulted in 637 upregulated and 98 downregulated genes. Figure 3d,e illustrate the DEG distribution that corresponds to the FC and p -adj. values. Appendix A1 in Data S1 provides details on the overall DEG results for both systems.

The combined set of differentially expressed genes in the AMBIC system (951 genes) and the BalanCD system (735 genes) were processed using GO enrichment analysis. Figure 4 shows the top 20 upregulated and the top 20 downregulated GO gene clusters for viral production (see Figures S1 and S2 for the top 100 GO gene clusters). Among the upregulated GO gene clusters shown in Figure 4a, more than half were related either to the innate immune response of the host cell (e.g., interferon signaling), different virus infection responses, or the host cell's antiviral responses. This reveals that AAV production stimulates the HEK293 cellular immune response. A few common metabolic processes (e.g., the glycolipid metabolic process, glycosaminoglycan metabolism) were downregulated and clustered in the GO enrichment (see Figure 4b). The next several sections elaborate on further analysis related to AAV production.

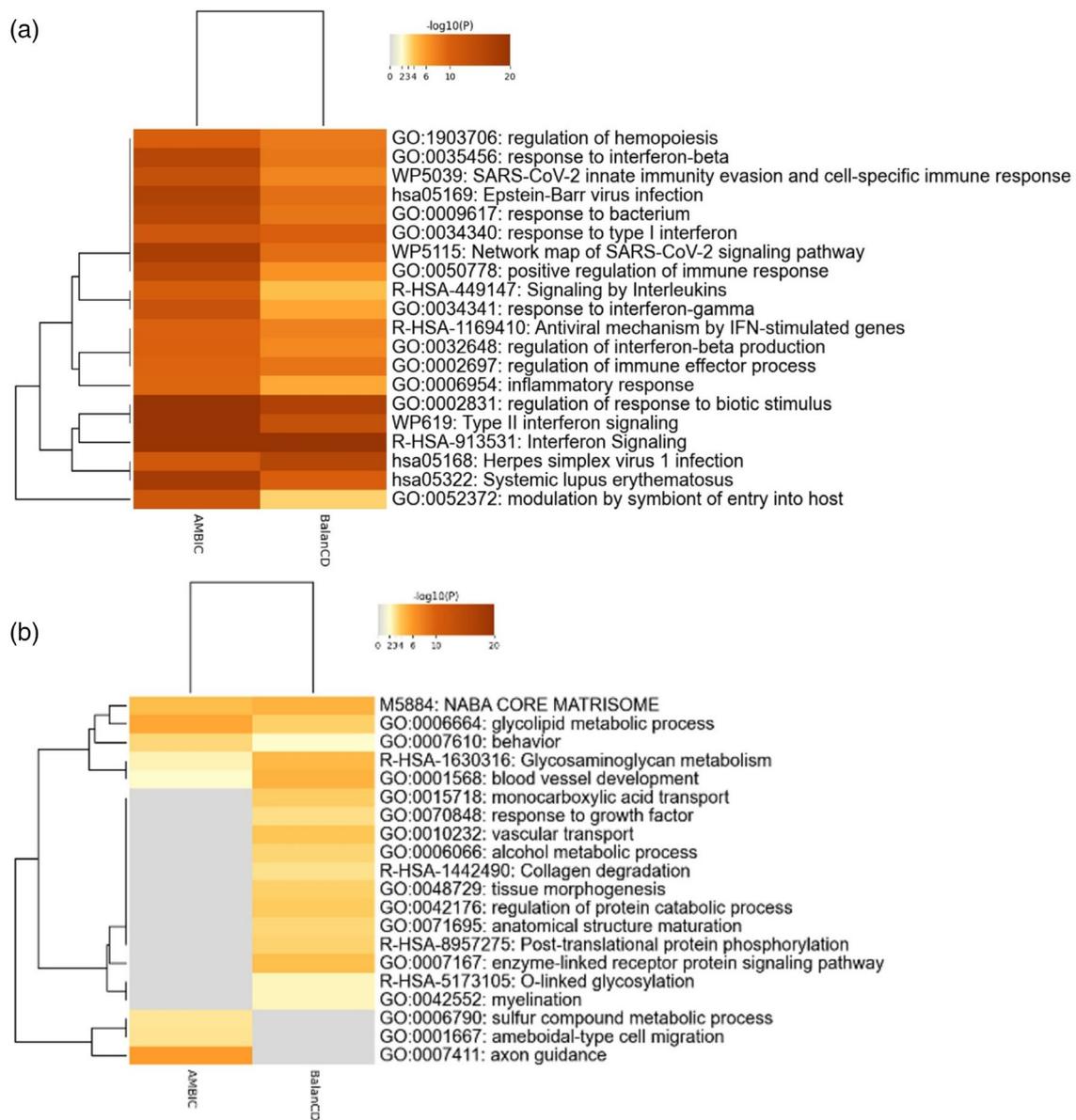


FIGURE 4 Enriched upregulated (a) and downregulated (b) clusters in the viral-producing states based on the gene ontology database for both AMBIC and BalanCD cell cultures. Differentially expressed genes in both systems with a 1.5-fold change threshold and p -adj. values <0.01 were processed for ontology analysis. The figure shows the comparison of the AMBIC and BalanCD cell cultures

3.3 | DEGs, functions, and pathway analysis

The transcriptomes of two viral producers were compared to their corresponding non-producing states on various post-transfection days to investigate the regulation of cell line-independent pathways for AAV viral production. The resulting list of DEGs with the threshold fold change (>1.5) and p -adj. value (<0.01) were enriched using the KEGG database (see Appendix A2 in Data S1 for the HEK-AMBIC system and A3 for the HEK-BalanCD system). The DEG and pathway analysis in this section focused on D2 (48 h post-transfection) because this was the most rapidly proliferating phase for AAV

production as shown in Figure 2b,d. Transcriptome changes relating to viral production that arise over time for viral production are discussed briefly in the Table S9. The primary enriched pathways include antiviral immunity signaling pathways and cellular stress-associated pathways for both cell lines. Significantly enriched pathways were further investigated to evaluate their correlation with viral production and their potentials to regulate productivity. Detailed statistical results on pathway enrichment can be found in Appendix A6 in Data S1 for the HEK-AMBIC system and Appendix A7 in Data S1 for the HEK-BalanCD system. The official full names of genes and proteins mentioned below can be found in the list of abbreviations.

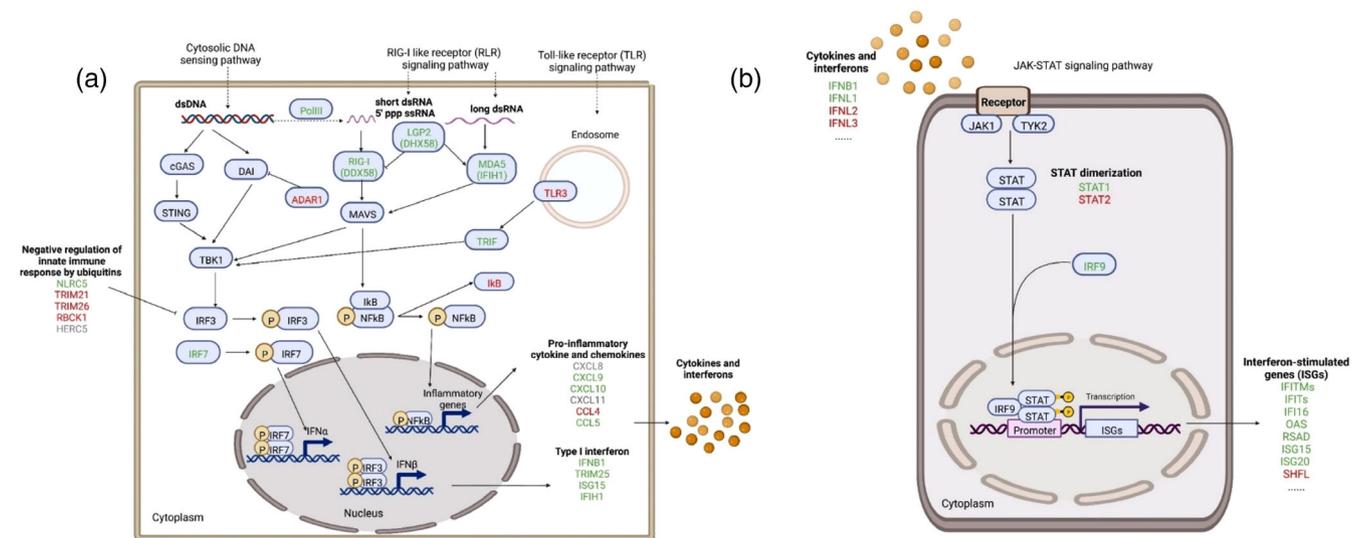


FIGURE 5 Innate immune response signaling pathways enriched in viral-producing states. (a) shows RIG-I like receptor (RLR) signaling, Toll-like receptor (TLR) signaling, and cytosolic DNA sensing pathways. (b) shows JAK-STAT signaling pathways. The genes shown in green were found to be common in both cell lines. Genes shown in red are specific to the HEK-AMBIC cell line. Genes shown in gray are specific to HEK-BalanCD cell line

3.4 | Innate immune responses

Several signaling pathways involved in the innate immune response, ranked as the top enriched KEGG pathways, were significantly upregulated in the viral-producing state. Host cell innate immunity often acts as the first line of defense against the spread of viruses. Once a pathogen-associated molecular pattern (e.g., viral nucleic acids) is recognized by the host's pathogen recognition receptor, a series of downstream signaling pathways is activated to stimulate the production of type I interferon (IFN), and the transcription of interferon-stimulated genes (ISGs).³² The expression of these genes results in the negative regulation of viral replication and the restriction of the viral production.³² The innate immune response signaling pathways shown in Figure 5 were enriched in the viral-producing state. These pathways include the *Toll-like receptor* (TLR) signaling pathway, the *retinoic acid-inducible gene-1-like receptor* (RLR) signaling pathway, and the downstream Janus kinase-signal transducer and activator of transcription (JAK-STAT) signaling pathway. Table S6 shows the FC and *p*-adj. values for these innate immune response signaling pathways in both cell lines. The JAK-STAT signaling pathway is downstream for all antiviral sensing pathways. The JAK-STAT signaling pathway describes all the commonly expressed genes in the downstream of the RLR, TLR, and cytosolic DNA sensing pathways (see Appendices A6 and A7 in Data S1 for the other enriched KEGG pathways).

3.4.1 | RLR signaling pathway

The RLR signaling pathway is the major sensing pathway for RNA viruses.³³ This study finds that this pathway was enriched for RAAV production (see Figure 5a). RIG-I (also referred to as DDX58) and

MDA5 (IFIH1) are receptor proteins for short dsRNA and long dsRNA, respectively.³³ Besides RNA source directly from evading viruses, the conversion of dsDNA to 5'ppp ssRNA via RNA polymerase III (PolIII) also allowed for the activation of the RLR signaling pathway by sensing cytosolic DNA.^{34,35} The overexpression of *DDX58*, *IFIH1*, and *POLR3C* (a subunit of *PolIII*) in both cell lines shown in the transcriptome data suggested the activation of the RLR signaling pathway. The detection of foreign nucleic acids facilitated the activation of the MAVS protein and the recruitment of the downstream transcription factor *TBK1*, triggering the phosphorylation of *IRF3* and *IRF7*. The transcriptome data indicate that *IRF7* in both the AMBIC (FC = 15.102, *p*-adj. = 0) and BalanCD (FC = 10.753, *p*-adj. = 2.92×10^{-45}) cell lines was upregulated in the viral-producing state. Then, phosphorylated *IRF7* translocated to the nucleus and triggered the production of type I IFNs and pro-inflammatory cytokines. The production of type I IFNs, pro-inflammatory cytokines, and chemokines was confirmed by the upregulated expression of multiple genes, including *IFNB1*, *TRIM25*, *CCL5*, and *CXCL9* (see Figure 5a).

3.4.2 | TLR signaling pathway

There are two categories of *Toll-like receptors* (TLR). One category consists of receptors on the cellular surface that detect viral proteins. The other category consists of receptors in the intracellular endosome that recognize viral nucleic acids. The transcriptome data demonstrate the enrichment of the TLR signaling pathway in both cell lines. Specifically, in the AMBIC cell line, *TLR3* (FC = 2.8, *p*-adj. = 3.74×10^{-7}), which is one type of endosome receptor, was found to be upregulated in the viral-producing state. This upregulation has been shown to be a host cell defense mechanism against viruses by limiting viral replication via

the type I IFN production.³⁶ The recognition of foreign components by TLR3 triggers the activation of *TRIF* (*TICAM1*). Enhanced *TRIF* expression was observed in both the BalanCD (FC = 10.232, p -adj. = 1.034×10^{-4}) and the AMBIC (FC = 1.547, p -adj. = 1.214×10^{-4}) cell lines. The activation of *TRIF* recruited the downstream transcription factor *TBK1* and phosphorylated *IRF7*. Similarly, the TLR signaling pathway can also produce type I IFN and cytokine release, which further triggers the downstream JAK-STAT signaling pathway.

3.4.3 | Cytosolic DNA sensing pathway

cGAS, together with *DAI*, are major cytosolic DNA sensors or receptor proteins that initiate the production of type I IFNs via phosphorylated *IRF3* or *IRF7*.³⁷ The DEG data and KEGG pathway analysis showed that the cytosolic DNA sensing-signaling pathway was enriched in both cell lines. The overexpression of *POLR3C*, *ADAR1*, and the production of interferons and proinflammatory cytokines, activating the downstream JAK-STAT signaling pathway, may result in this enrichment (see Figure 5a). Detailed fold changes and p -adj. value can be found in Table S6.

3.4.4 | JAK-STAT signaling pathway

Secreted IFNs bind to recognition receptors embedded in the cell surface, which triggers the activation of the JAK-STAT pathway.^{32,38} Significant upregulation of IFN expression (e.g., *IFNB1*, *IFNL1*) was observed in the viral-producing state in both cell lines. As shown in Figure 5b, phosphorylated STAT proteins formed heterotrimeric transcription factor complexes with *IRF9*. This was confirmed by the highly upregulated STAT protein and *IRF9* expression. The complex translocated to the nucleus, bound to the promoter, and initiated the expression of ISGs such as the *IFITM* family, the *IFIT* family, the *OAS* family, or *RSAD2*.

ISGs were shown to be indicators of the host cell's defense against viruses. The ISGs common to both cell lines are shown in green text in Figure 5b. These genes were induced by IFN production, and they have their own mechanisms to negatively regulate viral genome replication and thus limit the spread of the virus.³⁹⁻⁴⁵ Among the ISGs, the *IFITM* family of genes was found to be upregulated in the viral-producing condition, exhibiting an expression level five times higher than in the non-producing condition. These changes were more prominent in the AMBIC cell line (*IFITM1*, FC = 70.95, p -adj. = 0) than in the BalanCD cell line (*IFITM1*, FC = 10.265, p -adj. = 3.92×10^{-71}). These interferon-induced transmembrane proteins act as virus restriction factors via viral entry-dependent or viral entry-independent inhibition, such as suppressing viral protein synthesis or limiting viral replication.^{46,47} *IFIT5* is another interferon-induced binding protein involved in innate immune responses (FC = 5.098, p -adj. = 5.39×10^{-250} in the AMBIC cell line; FC = 2.448, p -adj. = 4.4×10^{-23} in the BalanCD cell line). This protein could inhibit viral replication through multiple mechanisms, such as restraining translation initiation, sequestering viral nucleic acids or proteins in the

cytoplasm, or binding to uncapped (or incompletely capped) viral RNA.⁴⁷ Table S6 provides more details on the FC and gene information of the ISGs involved in the enriched JAK-STAT pathway.

This significant enrichment in the host's immune response against viruses provides insight into the changes observed in cellular physiology that relate to rAAV production. This study proposes a few control strategies to modulate this response pathway. Virus replication is largely determined by two factors: the spread of virus replication and the induction of the antiviral state. Therefore, interferon-sensitive viruses can be difficult to cultivate to high titer in cells that produce IFNs.⁴⁸ In order to restrict IFN production and ISG expression, Stewart et al. proposed the use of inhibitors of the signaling proteins involved in IFN signaling pathways, such as BX795 (a *TBK1* inhibitor), *TPCA-1* (an *IKK2* inhibitor), or ruxolitinib (a *JAK1* inhibitor).⁴⁸ These small molecular inhibitors were utilized to block specific interferon pathways, and they were tested to successfully lower the innate immune response and enhance replication for Respiratory Syncytial Virus, Influenza Virus, and so on.⁴⁸ Similarly, the enrichment of the antiviral signaling pathway in rAAV production suggests the hypothesis that control of the IFN signaling pathway can contribute to enhanced viral production. The promising results in the literature on IFN inhibitors mentioned above supply a proof-of-concept that these IFN inhibitors can be used as media supplements to block or decrease the innate immune response pathways in HEK293 cells and enhance rAAV genome replication, thus improving cellular productivity.

Furthermore, the literature also reports that the critical genes triggered by the innate immune response (e.g., *RSAD2*,^{42,44,49} *OAS* family genes,⁴⁰ *IFIT* family genes),⁵⁰ affect virus genome replication. For example, the expression of the enzyme *RSAD2* could deplete cellular nucleotide pools and decrease the mitochondrial respiration rate via its radical activity, thereby restricting the genome replication of various RNA or DNA viruses.⁴⁴ Cell line engineering strategies (e.g., gene silencing, gene knockout) might be used to lower or eliminate the expression of those genes. However, the biological functions of these genes in terms of maintaining basic cellular activities must be explored further.

3.4.5 | Ubiquitin in the attenuation of the innate immune response

In addition to enriched innate immune response signaling pathways, some clustered genes function as negative regulators of the innate immune response via the ubiquitin modification system (see Figure 5a). FC, p -adj. values, and detailed biological functions of these genes can be found in Table S7. This ubiquitin modification strategy is often utilized by other viruses to evade antiviral signaling pathways and suppress IFN-I expression.⁵¹⁻⁵³ For example, *ICP0* is a viral protein expressed by the herpes simplex virus 1 (HSV1) with E3 ubiquitin ligase activity. *ICP0* has been shown to play an important role in minimizing the production of IFN and ISGs during HSV1 viral infections.⁵³ Specifically, *ICP0* targets activated *IRF3/IRF7* for degradation, inhibits

the activation of STAT proteins, and suppresses the expression of TRIMs.⁵³ Other than the result shown in this transcriptomic study, no literature has been reported about the ubiquitin modification system utilized by host cells for AAV production. One promising—but challenging—approach to enhancing virus productivity may be imitating other viruses' strategies to evade antiviral responses during the transfection process and eliminate the negative impacts of immune responses on viral genome replication. Co-transfection with pre-designed plasmids allows for the expression of foreign viral proteins (e.g., ICPO) in HEK293 cells, so they can serve their biological function and reduce the production of IFN and ISGs during rAAV production. However, we need to be aware that plasmids as foreign DNAs, could also result in toxicity and other stresses to host cells. Another approach to promoting efficient rAAV production can be amplifying innate-associated ubiquitin behaviors by overexpression, thus further restricting antiviral responses. Both methods are aimed at eliminating or lowering the viral components induced innate immune responses, thus benefiting for viral replication.

3.5 | Host cell stress responses

The AAV viral production cycle relies on the host cell's replication and translation machinery. Viral proteins and viral nucleic acids are foreign components that could trigger a series of host cell stress responses, such as autophagy, ER stress, or apoptosis. These were all found to be upregulated in the viral-producing states. The transcriptomic results regarding ER stress induced UPR are mainly discussed in this section. Table S7 provides details of the GO enrichment results and the FC values for the DEGs in both cell lines. Appendices A4 and A5 in Data S1 show more GO enrichment results.

ER stress stimulated by viral production and UPR: Viruses rely on host ER machinery for viral protein synthesis and modification. To ensure the cell's survival and viral replication, host cells have to respond to this stress and trigger the activation of unfolded protein response (UPR).⁵⁴ GO enrichment analysis shows the cellular response to ER stress and the activation of UPR in the viral-producing state. PERK-mediated UPR, serving as one major UPR pathway, was found closely related to AAV production in this study. In PERK-mediated UPR, the activation of PERK by the accumulation of unfolded or misfolded proteins, phosphorylates translation initiation factor EIF2 α , resulting in temporary protein synthesis attenuation. To ensure proper mRNA translation and protein synthesis, phosphorylated EIF2 α in host cells can be deactivated by downstream negative feedback signaling molecules, such as *EIF2AK2* and *PPP1R15A* (which encodes for GADD34).⁵⁵ It was observed that both *EIF2AK2* and *PPP1R15A* were upregulated for viral production in both cell lines. This demonstrates that viral production indeed induces ER stress, and host cells can resolve ER stress and restore protein synthesis by over-expressing negative feedback signaling genes such as GADD34. Furthermore, the overexpression of the heat shock protein HSPA6 was observed in both the AMBIC (FC = 4.376, p -adj. = 7.67×10^{-10}) and

the BalanCD (FC = 13.454, p -adj. = 4.94×10^{-13}) cell lines. In this case, stressful conditions induced the upregulation of HSPA6, which serves as an essential modulator to promote correct protein folding and ensure host cell survival in stressful environments.^{56,57}

To modulate the UPR, and to relieve ER stress, the small molecules that enhance the expression of the heat shock protein (i.e., HSPA6 activators) and inhibit the expression of the UPR receptor protein (i.e., PERK inhibitors) can be evaluated for their impact on viral productivity.^{58,59} Overexpression of genes involved in ER protein processing has been proved to achieve up to 97% volumetric titer improvement in the recombinant retroviral vector production.⁶⁰ Similar small molecule additives or cell line engineering approaches can be evaluated to resolve ER stress and recover AAV-related protein synthesis, thereby improving rAAV productivity. Additionally, ER stress and UPR pathways have also been reported for biologics production in Chinese Hamster Ovary (CHO) cells.^{61,62} Reported strategies might also offer another good starting point for relieving ER stress and therefore improving yield in the vector production: for example, chemical chaperones (e.g., trehalose, proline, glycerol) have been reported to increase recombinant protein productivity and decrease protein aggregation.^{63,64} Overexpression of genes encoding molecular chaperones also enhanced the expression of difficult-to-express recombinant proteins in HEK293 cell lines.^{65,66}

Apoptosis is another host stress response to defend against viral replication. From GO analysis, both cell lines exhibit the enrichment of the positive regulation of apoptotic process (GO:004305 and GO:0006915). See Appendix 4 in Data S1 for GO enrichment in HEK293-AMBIC and Appendix 5 in Data S1 in HEK293-BalanCD. The expression of *rep* and helper genes has been shown to induce cytotoxicity: Rep78 protein, the most toxic protein, was found to activate caspase-3 and induce cell apoptosis⁶⁷; E4orf4 was the major E4 product responsible for induction of p53-independent apoptosis.⁶⁸ Upon transfection, host cells were exposed not only to external pathogen components but also to transfection reagents (e.g., PEI pro). Within a relatively short 3-day cultivation period, viability for both cell lines was maintained above 80%. However, when compared to non-producing conditions, there was an obvious upregulation of apoptosis pathways induced by viral vector production (Appendices 4 and 5 in Data S1). In order to alleviate apoptosis-induced stress, small molecules such as antioxidants,⁶⁹ can be used as media supplements. Another potential strategy to facilitate rAAV production may be the overexpression of anti-apoptotic genes via cell engineering methods to delay apoptosis in the HEK293 cell line.¹⁸ A study has been confirmed the improvement of r-RV productivity via overexpression of B-cell lymphoma 2 (*Bcl2*) protein gene (anti-apoptotic).⁶⁰

Other GO-enriched cellular responses to stress were observed (e.g., the cellular response to glucose starvation, the response to the presence of reactive oxygen species, or enriched oxidant detoxification) to cope with such stressful conditions. FC, p -adj. value and biological functions of associated genes can be found in Table S7. More GO enrichment results can be found in Appendix 4 in Data S1 (HEK293-AMBIC) and 5 (HEK293-BalanCD).

3.6 | Metabolism

In both cell lines, *SLC43A2* was slightly overexpressed in the viral-producing state, as shown in Table S7. This encoded transmembrane protein was involved in transporting neutral amino acids. This overexpression could be related to the increased demand for the raw materials required for viral capsid synthesis in the viral-producing condition. More importantly, in the late stage of viral production, fatty acid metabolism was found to be downregulated in both cases (see Appendices A4 and A5 in Data S1). Furthermore, two genes involved in the *branched-chain amino acid* (BCAA) catabolic process, *ALDH6A1* and *ACADSB*, were found to be downregulated in the viral-producing state in both cell lines. BCAAs functioned as signaling molecules that regulate levels of glucose and lipids, the synthesis of proteins, and the immunity.⁷⁰ Further studies may be necessary to explore specific correlations between downregulated metabolism and viral production.

3.7 | Transcriptomic results validation by qPCR

The analysis and hypothesis elaborated above were based on the transcriptomic study. To validate the changes in the expression of main genes and enriched pathways for viral production, transcript levels of selected genes were further quantified using qPCR. Genes involved in mainly enriched pathways were chosen for the validation test. Since our transcriptomic analysis was mainly built upon the samples from post-transfection day 2 (D2), the most rapidly proliferating phase for rAAV production, D2 cDNA libraries were then utilized for validation experiments. Table S8 showed the normalized gene expression changes between viral producing and non-producing conditions for both HEK293-BalanCD and HEK293-AMBIC. Standard deviation (SD) and coefficient of variance (CV) were calculated based on the biological replicates.

Large upregulation of IFN- β , IRF7&9, and ISGs (*OAS1*, *RSAD2*, and *ISG15*) expression further confirmed the enrichment of innate immune responses. Overexpression of heat shock proteins (*HSPA6* and *PPP1R15A*) confirmed the activation of UPR and their beneficial functions of relieving the ER stress for viral vector production. qPCR results were comparable to those shown in the transcriptomic data. It cross-validated the transcriptomic results and analysis, and confirmed the common enriched pathways we identified for the parental HEK293 cell lines.

4 | CONCLUDING REMARKS

This study compares the transcriptomes of AAV-producing and non-producing groups over time using different sources of parental HEK293 cells adapted and cultured in their respective media. A transcriptomic variance was observed between the viral-producing and non-producing groups in both systems. Their transcriptomic features reveal pathways, including innate immune responses, cell stress

responses, and specific metabolisms that potentially impact rAAV production in parental HEK293 cells.

These transcriptomic results provide a mechanistic understanding to guide rational modifications (including cell line development and media optimization) to enhance rAAV production. The antiviral immune response is one of the most significant bottlenecks identified in viral production. This study proposes multiple strategies to inhibit this response: The first recommendation is to supplement the medium with small molecules that inhibit interferon-signaling proteins. This has the potential to effectively lower the production of interferon and eliminate its negative regulation of viral replication. Second, gene silencing or knockout methods can be attempted to restrict the expression of ISGs (e.g., *RSAD2*, the OAS family, the IFITM family) to eliminate their detrimental impact on viral production. Last, viral productivity may be modified by imitating the strategies used by other viruses to escape the innate immune response signaling pathways via co-transfection with plasmids that express specific viral proteins.

It has been reported that the host's cellular metabolism is reprogrammed during viral production and viral infection.^{71,72} However, our transcriptomic study provides limited insights into these changes in the host cell's metabolism. Future investigations should consider host cell metabolism for AAV production. It is also critical to understand the metabolic pathways related to viral production and the accumulation of inhibitory metabolites that restrict viral productivity.

AUTHOR CONTRIBUTIONS

Yongdan Wang: Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Qiang Fu:** Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Yong Suk Lee:** Data curation (supporting); investigation (supporting); methodology (supporting); writing – review and editing (supporting). **Sha Sha:** Conceptualization (lead); funding acquisition (lead); supervision (lead); writing – review and editing (supporting). **Seongkyu Yoon:** Funding acquisition (supporting); project administration (lead); supervision (equal); validation (equal); visualization (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

This work was funded and supported by the Advanced Mammalian Biomanufacturing Innovation Center (AMBIC) by means of the Industry-University Cooperative Research Center Program under the U.S. National Science Foundation (Grant numbers 1624684 and 2100075). The authors would like to express our gratitude to all AMBIC member companies for their mentorship and financial support. This work was conducted for a doctoral thesis, and it was partially funded by The National Institute for Innovation in Manufacturing Biopharmaceuticals (NIIMBL, Grant/Award Number: 70NANB17H002). The authors also appreciate Sartorius Data

Analytics for providing MODDE and SIMCA licenses for this study. The authors gratefully thank Dr. David McNally for the kind provision of HEK293 cell line. The authors also thank Jack Lepine from the UMass Lowell Core Research Facility (CRF) for his kind support.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/btpr.3346>.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are provided in the supplemental files.

ORCID

Seongkyu Yoon  <https://orcid.org/0000-0002-5330-8784>

REFERENCES

- Ferrari FK, Xiao X, Mccarty D, Samulski RJ. New developments in the generation of ad-free, high-titer rAAV gene therapy vectors. *Nat Med*. 1997;3(11):1295-1297.
- Athanasopoulos T, Fabb S, Dickson G. Gene therapy vectors based on adeno-associated virus: characteristics and applications to acquired and inherited diseases. *Int J Mol Med*. 2000;6(4):363-438.
- Patel A, Zhao J, Duan D, Lai Y. *Adeno-Associated Virus Vectors*. Springer; 2019:19-33.
- Clément N, Grieger JC. Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Mol Ther-Meth Clin Dev*. 2016;3:16002.
- Weitzman MD, Linden RM. *Adeno-Associated Virus*. Springer; 2012: 1-23.
- Meier AF, Fraefel C, Seyffert M. The interplay between adeno-associated virus and its helper viruses. *Viruses*. 2020;12(6):662.
- Sha S, Maloney AJ, Katsikis G, et al. Cellular pathways of recombinant adeno-associated virus production for gene therapy. *Biotechnol Adv*. 2021;49:107764.
- Wang D, Tai PWL, Gao G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat Rev Drug Discov*. 2019;18(5):358-378. doi:10.1038/s41573-019-0012-9
- Grieger JC, Soltys SM, Samulski RJ. Production of recombinant adeno-associated virus vectors using suspension HEK293 cells and continuous harvest of vector from the culture media for GMP FIX and FLT1 clinical vector. *Mol Ther*. 2016;24(2):287-297. doi:10.1038/mt.2015.187
- Bingnan Gu VB, Dong W, Pham H, et al. Establishment of a scalable manufacturing platform for in-silico-derived ancestral adeno-associated virus vectors. *Cell Gene Ther Insights*. 2018;4:741-751. doi:10.18609/cgti.2018.078
- Zhao H, Lee K-J, Daris M, et al. Creation of a high-yield AAV vector production platform in suspension cells using a Design of Experiment Approach. *Mol Ther-Meth Clin Dev*. 2020;18:312-320.
- Nguyen TN, Sha S, Hong MS, et al. Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. *Mol Ther-Meth Clin Dev*. 2021;21:642-655.
- Sha S, Bhatia H, Yoon S. An RNA-seq based transcriptomic investigation into the productivity and growth variants with Chinese hamster ovary cells. *J Biotechnol*. 2018;271:37-46.
- Kang S, Ren D, Xiao G, et al. Cell line profiling to improve monoclonal antibody production. *Biotechnol Bioeng*. 2014;111(4):748-760.
- Rodrigues A, Formas-Oliveira A, Bandeira V, Alves P, Hu W, Coroadinha A. Metabolic pathways recruited in the production of a recombinant enveloped virus: mining targets for process and cell engineering. *Metab Eng*. 2013;20:131-145.
- Futschik ME, Morkel M, Schäfer R, Sers C. The human transcriptome: implications for understanding, diagnosing, and treating human disease. *Mol Pathol*. 2018;135-164.
- Lin Y-C, Boone M, Meuris L, et al. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat Commun*. 2014;5(1):1-12.
- Abaandou L, Quan D, Shiloach JJC. Affecting HEK293 cell growth and production performance by modifying the expression of specific genes. *Cells*. 2021;10(7):1667.
- Ye Q, Phan T, Hu W-S, et al. Transcriptomic characterization reveals attributes of high influenza virus productivity in MDCK cells. 2021; 13(11):2200.
- Bell TA, Velappan N, Gleasner CD, et al. Nonclassical autophagy activation pathways are essential for production of infectious influenza a virus in vitro. 2022;117(2):508-524.
- Yu G, Lin Y, Tang Y, Diao YJljoms. *Comparative transcriptomic analysis of immune-related gene expression in duck embryo fibroblasts following duck tembusu virus infection*. 2018;19(8):2328.
- Almuqrin A, Davidson AD, Williamson MK, et al. SARS-CoV-2 vaccine ChAdOx1 nCoV-19 infection of human cell lines reveals low levels of viral backbone gene transcription alongside very high levels of SARS-CoV-2 S glycoprotein gene transcription. 2021;13(1):1-17.
- Havranek KE, White LA, Lanchy J-M, Lodmell JSJPO. Transcriptome profiling in rift valley fever virus infected cells reveals modified transcriptional and alternative splicing programs. 2019;14(5):e0217497.
- Bolger AM, Lohse M, Usadel BJB. Trimmomatic: a flexible trimmer for illumina sequence data. 2014;30(15):2114-2120.
- Langmead B, Salzberg SLJN. Fast gapped-read alignment with bowtie 2. 2012;9(4):357-359.
- Li B, Dewey CNJB. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. 2011;12(1):1-16.
- Kucukural A, Yukselen O, Ozata DM, Moore MJ, Garber MJB. DEBrowser: interactive differential expression analysis and visualization tool for count data. 2019;20(1):1-12.
- Love MI, Huber W, Anders SJG. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. 2014;15(12):1-21.
- Benjamini Y, Hochberg YJJ. Controlling the false discovery rate: a practical and powerful approach to multiple testing. 1995;57(1): 289-300.
- Huang DW, Sherman BT, Tan Q, et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. 2007;35(suppl_2):W169-W175.
- Ruiz-Perez D, Guan H, Madhivanan P, Mathee K, Narasimhan G. So you think you can PLS-DA? *BMC Bioinformatics*. 2020;21(1):1-10.
- Raftery N, Stevenson NJ. Advances in anti-viral immune defence: revealing the importance of the IFN JAK/STAT pathway. *Cell Mol Life Sci*. 2017;74(14):2525-2535.
- Brisse M, Ly H. Comparative structure and function analysis of the RIG-I-like receptors: RIG-I and MDA5. *Front Immunol*. 2019;10:1586.
- Jia J, Fu J, Tang H. Activation and evasion of RLR signaling by DNA virus infection. *Front Microbiol*. 2021;12:1-9.
- Chiu Y-H, MacMillan JB, Chen ZJ. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell*. 2009;138(3):576-591.
- Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on toll-like receptors. *Nat Immunol*. 2010;11(5): 373-384.

37. Lin Y, Zheng C. A tug of war: DNA-sensing antiviral innate immunity and herpes simplex virus type 1 infection. *Front Microbiol.* 2019;10:2627.
38. Sadler AJ, Williams BR. Interferon-inducible antiviral effectors. *Nat Rev Immunol.* 2008;8(7):559-568.
39. Jønsson K, Laustsen A, Krapp C, et al. IFI16 is required for DNA sensing in human macrophages by promoting production and function of cGAMP. *Nat Commun.* 2017;8(1):1-17.
40. Drappier M, Michiels T. Inhibition of the OAS/RNase L pathway by viruses. *Curr Opin Virol.* 2015;15:19-26.
41. Gusho E, Baskar D, Banerjee S. New advances in our understanding of the "unique" RNase L in host pathogen interaction and immune signaling. *Cytokine.* 2020;133:153847.
42. Honarmand Ebrahimi K, Vowles J, Browne C, McCullagh J, James WS. ddhCTP produced by the radical-SAM activity of RSAD2 (viperin) inhibits the NAD⁺-dependent activity of enzymes to modulate metabolism. *FEBS Lett.* 2020;594(10):1631-1644.
43. Perng Y-C, Lenschow DJ. ISG15 in antiviral immunity and beyond. *Nat Rev Microbiol.* 2018;16(7):423-439.
44. Ebrahimi KH, Howie D, Rowbotham JS, McCullagh J, Armstrong FA, James WS. Viperin, through its radical-SAM activity, depletes cellular nucleotide pools and interferes with mitochondrial metabolism to inhibit viral replication. *FEBS Lett.* 2020;594(10):1624-1630.
45. Wu Y, Yang X, Yao Z, et al. C19orf66 interrupts Zika virus replication by inducing lysosomal degradation of viral NS3. *PLoS Negl Trop Dis.* 2020;14(3):e0008083.
46. Yáñez DC, Ross S, Crompton T. The IFITM protein family in adaptive immunity. *Immunology.* 2020;159(4):365-372.
47. Diamond MS, Farzan M. The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nat Rev Immunol.* 2013;13(1):46-57.
48. Stewart CE, Randall RE, Adamson CS. Inhibitors of the interferon response enhance virus replication in vitro. *PLoS One.* 2014;9(11):e112014.
49. Dumbrepatil AB, Ghosh S, Zegalia KA, et al. Viperin interacts with the kinase IRAK1 and the E3 ubiquitin ligase TRAF6, coupling innate immune signaling to antiviral ribonucleotide synthesis. *J Biol Chem.* 2019;294(17):6888-6898.
50. Zhou X, Michal JJ, Zhang L, et al. Interferon induced IFIT family genes in host antiviral defense. 2013;9(2):200.
51. Heaton SM, Borg NA, Dixit VM. Ubiquitin in the activation and attenuation of innate antiviral immunity. *J Exp Med.* 2016;213(1):1-13.
52. Viswanathan K, Früh K, DeFilippis V. Viral hijacking of the host ubiquitin system to evade interferon responses. *Curr Opin Microbiol.* 2010;13(4):517-523.
53. Zhu H, Zheng CJM, Reviews MB. The race between host antiviral innate immunity and the immune evasion strategies of herpes simplex virus 1. 2020;84(4):e00099.
54. He B. Viruses, endoplasmic reticulum stress, and interferon responses. *Cell Death Diff.* 2006;13(3):393-403.
55. Hetz C, Zhang K, Kaufman RJ. Mechanisms, regulation and functions of the unfolded protein response. *Nat Rev Mol Cell Biol.* 2020;21(8):421-438.
56. Su Y-S, Hwang L-H, Chen C-J. Heat shock protein A6, a novel HSP70, is induced during enterovirus A71 infection to facilitate internal ribosomal entry site-mediated translation. *Front Microbiol.* 2021;12:664955.
57. Iyer K, Chand K, Mitra A, Trivedi J, Mitra D. Diversity in heat shock protein families: functional implications in virus infection with a comprehensive insight of their role in the HIV-1 life cycle. *Cell Stress Chaperones.* 2021;26(5):743-768.
58. Calamini B, Silva MC, Madoux F, et al. ML346: a novel modulator of proteostasis for protein conformational diseases. *Probe Reports from the NIH Molecular Libraries Program [Internet].* 2013.
59. Smith AL, Andrews KL, Beckmann H, et al. Discovery of 1 h-pyrazol-3 (2 h)-ones as potent and selective inhibitors of protein kinase r-like endoplasmic reticulum kinase (PERK). *J Med Chem.* 2015;58(3):1426-1441.
60. Formas-Oliveira AS, Basílio JS, Rodrigues AF, Coroadinha AS. Overexpression of ER protein processing and apoptosis regulator genes in human embryonic kidney 293 cells improves gene therapy vectors production. *Biotechnol J.* 2020;15(9):1900562.
61. Omasa T, Takami T, Ohya T, et al. Overexpression of GADD34 enhances production of recombinant human antithrombin III in Chinese hamster ovary cells. *J Biosci Bioeng.* 2008;106(6):568-573.
62. Prashad K, Mehra S. Dynamics of unfolded protein response in recombinant CHO cells. *Cytotechnology.* 2015;67:237-254.
63. Onitsuka M, Tatsuzawa M, Asano R, et al. Trehalose suppresses antibody aggregation during the culture of Chinese hamster ovary cells. 2014;117(5):632-638.
64. Hwang SJ, Jeon CJ, Cho SM, Lee GM, SKJbP Y. Effect of chemical chaperone addition on production and aggregation of recombinant flag-tagged COMP-angiopoietin 1 in chinese hamster ovary cells. 2011;27(2):587-591.
65. Gong Y, Jiang JH, Li ST. Functional expression of human $\alpha 7$ nicotinic acetylcholine receptor in human embryonic kidney 293 cells. *Mol Med Rep.* 2016;14(3):2257-2263.
66. Cain K, Peters S, Hailu H, et al. A CHO cell line engineered to express XBP1 and ERO1- α has increased levels of transient protein expression. *Biotechnol Prog.* 2013;29(3):697-706.
67. Schmidt M, Afione S, Kotin RM. Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53. *J Virol.* 2000;74(20):9441-9450.
68. Marcellus RC, JeN L, Boivin D, Shore GC, Ketner G, Branton PE. The early region 4 orf4 protein of human adenovirus type 5 induces p53-independent cell death by apoptosis. *J Virol.* 1998;72(9):7144-7153.
69. Fernández-Frías I, Pérez-Luz S, Díaz-Nido JJMT. Enhanced production of herpes simplex virus 1 amplicon vectors by gene modification and optimization of packaging cell growth medium. 2020;17:491-496.
70. Nie C, He T, Zhang W, Zhang G, Ma X. Branched chain amino acids: beyond nutrition metabolism. *Int J Mol Sci.* 2018;19(4):954.
71. Thaker SK, Ch'ng J, Christofk HR. Viral hijacking of cellular metabolism. *BMC Biol.* 2019;17(1):1-15.
72. Prusinkiewicz MA, Mymryk JS. Metabolic reprogramming of the host cell by human adenovirus infection. *Viruses.* 2019;11(2):141.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wang Y, Fu Q, Lee YS, Sha S, Yoon S. Transcriptomic features reveal molecular signatures associated with recombinant adeno-associated virus production in HEK293 cells. *Biotechnol. Prog.* 2023;39(4):e3346. doi:10.1002/btpr.3346