

# Production of rAAV by plasmid transfection induces antiviral and inflammatory responses in suspension HEK293 cells

Cheng-Han Chung,<sup>1</sup> Christopher M. Murphy,<sup>1</sup> Vincent P. Wingate,<sup>1</sup> Jeffrey W. Pavlicek,<sup>1</sup> Reiko Nakashima,<sup>3</sup> Wei Wei,<sup>1</sup> Douglas McCarty,<sup>2</sup> Joseph Rabinowitz,<sup>2</sup> and Erik Barton<sup>1</sup>

<sup>1</sup>Pfizer Inc., Worldwide Research, Development and Medical, Bioprocess Research and Development, Morrisville, NC 27560, USA; <sup>2</sup>Pfizer Inc., Worldwide Research, Development and Medical, Rare Disease Research Unit, Morrisville, NC 27560, USA; <sup>3</sup>Pfizer Inc., Worldwide Research, Development and Medical, Simulation and Modeling Sciences, Cambridge, MA 02139, USA

**Recombinant adeno-associated virus (rAAV) is a clinically proven viral vector for delivery of therapeutic genes to treat rare diseases. Improving rAAV manufacturing productivity and vector quality is necessary to meet clinical and commercial demand. These goals will require an improved understanding of the cellular response to rAAV production, which is poorly defined. We interrogated the kinetic transcriptional response of HEK293 cells to rAAV production following transient plasmid transfection, under manufacturing-relevant conditions, using RNA-seq. Time-series analyses identified a robust cellular response to transfection and rAAV production, with 1,850 transcripts differentially expressed. Gene Ontology analysis determined upregulated pathways, including inflammatory and antiviral responses, with several interferon-stimulated cytokines and chemokines being upregulated at the protein level. Literature-based pathway prediction implicated multiple pathogen pattern sensors and signal transducers in up-regulation of inflammatory and antiviral responses in response to transfection and rAAV replication. Systematic analysis of the cellular transcriptional response to rAAV production indicates that host cells actively sense vector manufacture as an infectious insult. This dataset may therefore illuminate genes and pathways that influence rAAV production, thereby enabling the rational design of next-generation manufacturing platforms to support safe, effective, and affordable AAV-based gene therapies.**

## INTRODUCTION

Adeno-associated virus (AAV) belongs to the *Dependoparvovirus* genus of the *Parvoviridae* family, as AAV-infected cells become permissive for viral replication only with the presence of helper viruses, including *Adenoviridae* or *Herpesviridae*. In the absence of helper virus, AAV establishes a latent infection in transduced cells and is not associated with known disease in humans. The only AAV sequences required for persistence of the viral genome in the cell are the short inverted terminal repeats (ITRs). When a transgene expression cassette is flanked by ITRs, this cassette can be packaged in

AAV capsids provided in *trans* in producer cells, and the resulting recombinant AAV (rAAV) vector can be used to transduce animal or mammalian cells and tissues, leading to stable episomal maintenance and expression of the transgene. The stability and safety of rAAV transduction have been leveraged as a clinically proven vector for gene therapy of monogenic rare diseases, by delivery of a genome encoding a therapeutic transgene expression cassette. Approval of the first rAAV gene therapies by regulatory agencies has led to a sharp increase in the number of clinical trials using these vectors. Depending upon the indication, high doses of vector, often exceeding  $10^{15}$  vector genomes (vg) per patient, may be required.<sup>1</sup> These factors have greatly challenged current manufacturing capability in the field, and therefore increased productivity is critical to meet expanding clinical and predicted commercial vector demand.

Because of the ease of culture and transfection, the human cell line HEK293 is commonly used to produce rAAV via transient transfection with plasmids that provide the required components, including adenoviral helper genes, the therapeutic transgene cassette, the AAV replication machinery, and structural proteins.<sup>2</sup> Furthermore, HEK293 stably express adenovirus E1A and E1B, which was the result of the transformation of this line following transfection with adenovirus type 5 genomic DNA.<sup>3–5</sup> Both E1A and E1B are required for effective AAV replication. To increase manufacturability of rAAV at commercial scales, HEK293 cells have been adapted to growth in suspension (in the absence of serum) while maintaining high transfection efficiency.<sup>6–9</sup> Despite these process advances, volumetric yields of rAAV are still critically limiting to broader use of rAAV as a cost-effective gene therapy vector. In addition, there is an increasing focus on improving the quality of rAAV vectors, primarily by improving the efficiency and fidelity of vector genome packaging.<sup>10</sup>

Received 9 November 2022; accepted 13 January 2023;  
<https://doi.org/10.1016/j.omtm.2023.01.002>.

**Correspondence:** Erik Barton, Pfizer Inc., Worldwide Research, Development and Medical, Bioprocess Research and Development, Morrisville, NC 27560, USA.

**E-mail:** [erik.s.barton@pfizer.com](mailto:erik.s.barton@pfizer.com)

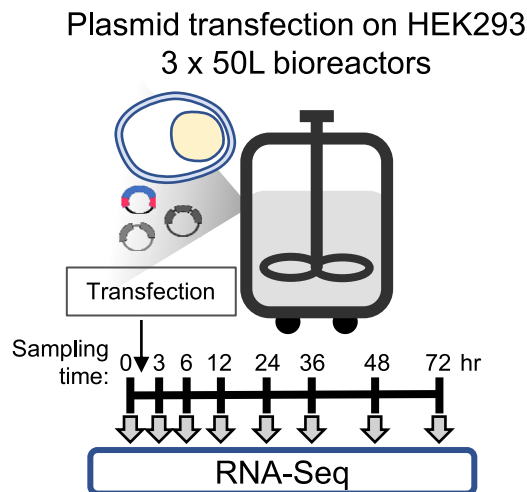
Achieving these targets of improved manufacturing yield and vector quality will require focused investment in understanding the molecular mechanisms of rAAV replication in manufacturing-relevant systems. rAAV production requires multiple poorly understood virus-host cell interactions, and the host cell response to these processes has not been defined. Human cells have conserved intrinsic sensors of, and defense pathways against, exogenous DNA and virus replication. Thus, initiation of rAAV production by plasmid transfection presents multiple stresses that have the potential to induce cellular responses, and these may in turn influence productivity. These potential stresses include cellular damage due to transfection compounds, the introduction of bacteria-derived DNA, residual endotoxin, the activities of adenoviral and AAV gene products, and ongoing AAV DNA replication. Previous studies found that transfection with polyethylenimine (PEI), which is commonly used for commercial rAAV production, triggers an inflammatory response, characterized by protein kinase C (PKC) activation and cytotoxicity that is cell type dependent.<sup>11-13</sup>

Expression of rAAV replication components, including viral genomes and adenovirus helper viral genes, activates numerous stress response pathways. For example, the large AAV replicase protein (Rep78) induces cell-cycle arrest by sequestering Cdc25A and nicking of cellular DNA.<sup>14,15</sup> The Toll-like receptor (TLR) pathogen sensor pathway TLR9-MyD88 is triggered by unmethylated CpG dinucleotides in AAV genomes during transduction, but the role of endosomal pathogen sensors in rAAV production has not been explored.<sup>16-18</sup> Helper gene E4 induces a prolonged DNA damage response by inhibiting protein phosphatase 2A and the MRE11-RAD50-NBS1 (MRN) complex, leading to initiation of cell death.<sup>19,20</sup> Pattern recognition receptors (PRRs) that sense bacterial or viral components are known to induce pathways that activate antiviral function, with the potential to inhibit rAAV production. HEK293 cells are generally thought to express low levels of sensors known for exogenous stimuli detection including cyclic GMP-AMP synthetase (cGAS), stimulator of interferon gene (STING), Toll-like receptors, and MDA5 proteins, which may render HEK293 easily transfectable.<sup>21-23</sup> However, it is not known whether rAAV production results in a coordinated cellular response that can alter rAAV productivity or vector quality.

Here we interrogated cellular responses to rAAV production by sequencing the transcriptome of a high-density, manufacturing-scale, transfected suspension HEK293 culture. Canonical pathway analysis as well as time-series analysis were conducted to identify host genes and pathways that respond to, and therefore may affect, rAAV production. We found enrichment for pathways involved in antiviral response and confirmed protein induction of antiviral and anti-pathogen inflammatory response genes. This systematic analysis not only presents an overview of the cellular pathways that respond to rAAV production but also sheds light on potential modulators of rAAV yield and vector quality.

## RESULTS

To understand the response of host cells to rAAV vector production via plasmid transfection, samples were collected throughout the course of three 50 L production batches and analyzed using bulk



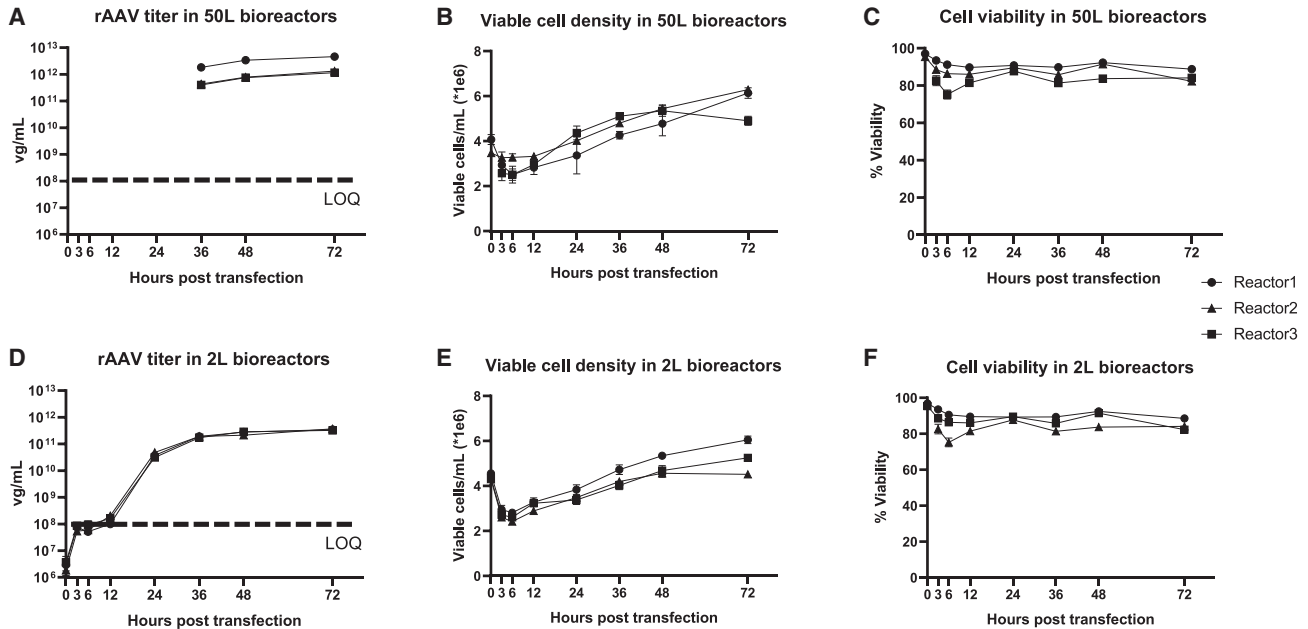
**Figure 1. The experimental design and workflow of transcriptome analysis of rAAV-producing HEK293 cells in production bioreactors**

Three independent 50 L single-use bioreactors (SUBs) were each sampled prior to transfection and at 7 time points post-transfection. RNA was prepared from these samples and subjected to RNA-seq.

RNA sequencing (RNA-seq) of polyadenylated transcripts. The rAAV produced in these lots encoded a transgene designed to treat Duchenne muscular dystrophy (DMD) delivered by the AAV9 capsid, and manufacture was enabled by a helper plasmid encoding the adenovirus E2A, E4, and VA RNA genes.<sup>9</sup> Samples were collected immediately prior to transfection (0 h) and at 3, 6, 12, 24, 36, 48, and 72 h post-transfection (hpt) (Figure 1). To validate the RNA-seq results via gene-specific assays and to define rAAV replication kinetics, three 2 L batches were also conducted under conditions that have been qualified to replicate the 50 L process (Figure 2). The titer of encapsidated rAAV genomes showed that the packaged vectors were predominantly produced between 12 and 24 hpt but continued to increase before harvest (Figures 2A and 2D).

RNA-seq quantification identified 17,209 expressed transcripts in at least one of the 24 samples (Figure S1). The result of principal-component analysis (PCA) on all expressed transcripts showed clear clustering by sampling time (Figure S2), indicating a consistent response to rAAV production via plasmid transfection among biological replicates. Furthermore, the time-series transcriptomic samples were ordered in relation to sampling times on the PCA graph as commonly seen with longitudinal samples. The results from PCA, combined with similar cell growth and viability among replicates (Figure 2), suggested that the transcriptomic regulation in transfected cells was highly reproducible between replicates.

To characterize the cellular transcriptional response to rAAV manufacture, we identified differentially expressed genes (DEGs) in all post-transfection time points relative to pre-transfection cells (0 h). In total, 1,850 unique DEGs were identified across all time points (Figure S1B). The number of upregulated genes increased over time



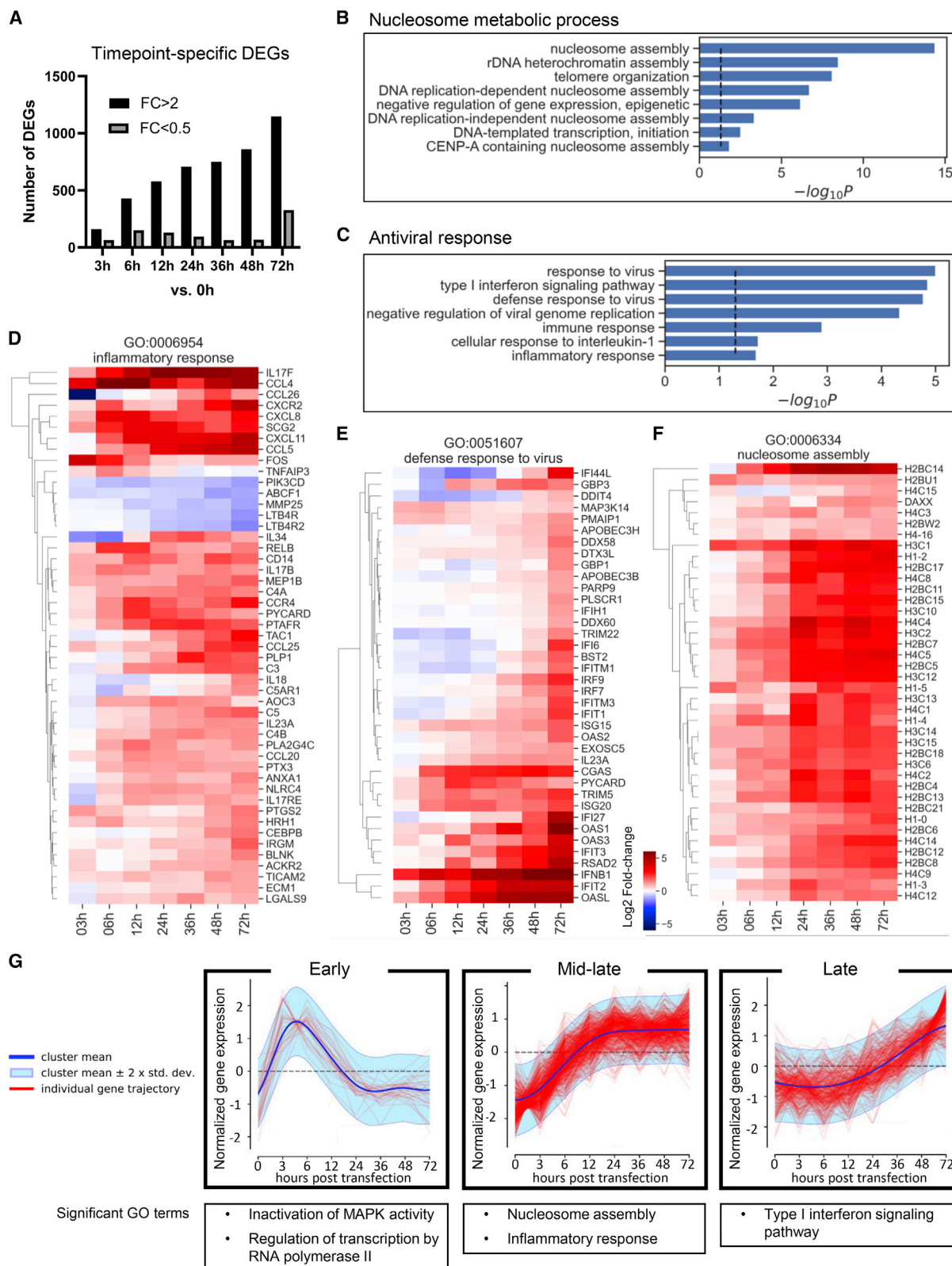
**Figure 2. Kinetics of rAAV production and cell growth in 50 L and 2 L bioreactors**

(A–C) For each time point selected for transcriptome analysis, the vector genome (vg) titer, viable cell density (VCD) and percentage viable cells were determined for each of the three 50 L bioreactors. (D–F) An additional set of three independent 2 L bioreactors were performed and sampled in the same manner, and vg titer, VCD, and percentage viability were determined. Viral genome titer was determined by rAAV qPCR described in [materials and methods](#). VCD and cell viability were measured using a Vi-Cell XR cell viability analyzer (Beckman Coulter Life Sciences). Dashed lines in (A) and (D) indicate the limit of quantification (LOQ) for AAV titer using the qPCR method. Error bars indicate SD of measurements among technical replicates.

post-transfection (Figure 3A). To better understand how host cells were responding to transfection and rAAV vector production, we performed Gene Ontology (GO) enrichment analysis to identify altered pathways. Using 1,850 DEGs in Gene Ontology analysis identified significant pathways involved in DNA/chromatin organization and DNA metabolic processes (Figure 3B). Interestingly, most of the DEGs involved in these pathways were histone subunits induced 24 h post-transfection, by which time rAAV production was fully underway (Figure 3F). Another subset of significant pathways was related to the innate immune response to virus (Figure 3C). Many interferon-stimulated genes (ISGs) were significantly enriched (Figure 3E). Expression levels of ISGs generally increased until the time of harvest (72 h post-transfection). The Gene Ontology category Inflammatory Response was also enriched in the DEG set (Figure 3D). A subset of chemokines and chemokine receptors was induced by 12 h after transfection. Key components of the NLRC4 inflammasome including NLRC4 and PYCARD (also known as ASC) involved in the activation of inflammatory and apoptotic pathways peaked at 12 h post-transfection. IL-17F, a pro-inflammatory cytokine that activates expression of other chemokines, was induced between the 24 and 36 h time points.<sup>24</sup> These DEGs identified in gene ontologies related to antiviral response suggest an active response to pathogen-associated stimuli in transfected HEK293.

Although pathway analysis can discern the general functional categories of gene expression in response to rAAV production, it does

not identify the triggers of these responses. To begin to dissect this mechanistic question, DEGs were clustered on the basis of their expression pattern via time-series analysis followed by Gene Ontology enrichment analysis. This time-series analysis allows transcriptional changes to be correlated with the processes required for transfection, rAAV replication, and vector assembly. A pattern clustering method on the basis of Dirichlet process Gaussian process (DPGP) was used to categorize similar expression dynamics on time-series data without a pre-specified number of clusters.<sup>25</sup> By this method we identified 13 clusters, with 142 genes on average per cluster (Figure S5). Clusters containing at least 40 genes were then subjected to Gene Ontology enrichment analysis, to avoid the high risk for false-positive enrichment associated with small gene sets.<sup>26</sup> Five of eight clusters had significantly enriched biological processes using Gene Ontology enrichment analysis (false discovery rate [FDR] <0.05) (Figure S6). These five clusters covered 62.8% of the DEGs. On the basis of the time that the normalized transcription reached above the  $y = 0$  axis, clusters with significantly enriched gene ontologies were categorized into early, mid-late and late stages. Genes in the early cluster were induced at 3 and 6 hpt and reverted to basal expression level by 12 hpt (Figure 3G). Notable genes in this cluster included dual-specificity phosphatases (DUSPs) 1, 8, and 16, which regulate mitogen-activated protein kinase (MAPK) activity. (Figures 3G and S6). The MAPK signaling pathways participate in transcription of immediate-early genes that respond to extracellular stimuli, including early growth response 1 (EGR1), FOS, JUN, and



(legend on next page)



MYC (Figure S6).<sup>27</sup> The DUSPs feed back to negatively regulate stress kinase activity and are MAPK signaling-inducible genes responding to acute environmental stimuli such as transfection and pathogen recognition receptor signaling.<sup>28</sup> The protein expression analysis of EGR1 validated the kinetic of EGR1 transcription across time points (Figure 4A). On the basis of the immediate-early timing and MAPK pathway involvement, these genes are potentially induced by transfection-associated stress or the earliest stages of rAAV and adenovirus helper gene expression.<sup>29</sup>

Genes in clusters that showed upregulated transcription in the mid-late and late clusters constituted the majority DEGs and consisted primarily of nucleosome assembly, inflammatory response, and type I interferon-driven pathways described above (Figures 3B–3G). The mid-late cluster included inflammasome-regulated genes and histone subunits and was characterized by significant up-regulation 6 h post-transfection and sustained up-regulation 12 h post-transfection (Figure 3G). The late responding cluster included DEGs that were not induced until 24 h post-transfection. Among these, type I interferon induced genes were significantly enriched. Expression of interferon beta (IFNB1) was detected at these time points, albeit at low levels, and correlated with up-regulation of multiple ISGs, including IFIT2 and OASL. This expression pattern was confirmed by qRT-PCR in samples from transfected 2 L bioreactors, which showed a consistent trend of mature polyadenylated IFNB1, IFIT2, and OASL mRNAs (Figures 4B, 4E, and 4F). Interestingly, however, interferon beta protein could not be detected in culture media by ELISA at any time point after transfection. This may be due to rapid binding to ubiquitously expressed IFN receptors on HEK cells, maintaining steady-state levels of IFNB below the limit of detection. CCL5, a pro-inflammatory ISG that can also be induced by activation of NF- $\kappa$ B and IRF1,<sup>30,31</sup> was strongly upregulated in the late cluster. CCL5 transcription was upregulated as early as 6 h post-transfection, while CCL5 secretion was significantly increased 24 h post-transfection (Figures 4C and 4H). Transcription of another ISG cytokine, IL-8, was also significantly increased followed by increase of IL-8 protein in media (Figures 4D and 4I).<sup>32</sup> Note that absolute AAV titers are not directly comparable between 50 L single-use bioreactors (SUBs) and 2 L bioreactors because of different qPCR primers and probes used to detect packaged AAV genomes (described in materials and methods). Nevertheless, the expression profiles of validated genes between RNA-seq in samples from 50 L SUBs and qRT-PCR in samples from 2 L biore-

actor were still highly aligned, indicating comparable cellular responses among scaling of bioreactors. These results suggested that transcriptional up-regulation of innate immune response pathways was induced as early as 12 h post-transfection, with low level expression of type I interferons, which in turn induced cytokine production and the up-regulation of ISGs later in the process. Up-regulation of innate immune response in the cells (Figure 2D) correlated with the exponential increase in rAAV titer between 12 and 24 hpt, suggesting that HEK293 cells detect viral replication and respond with antiviral gene expression.

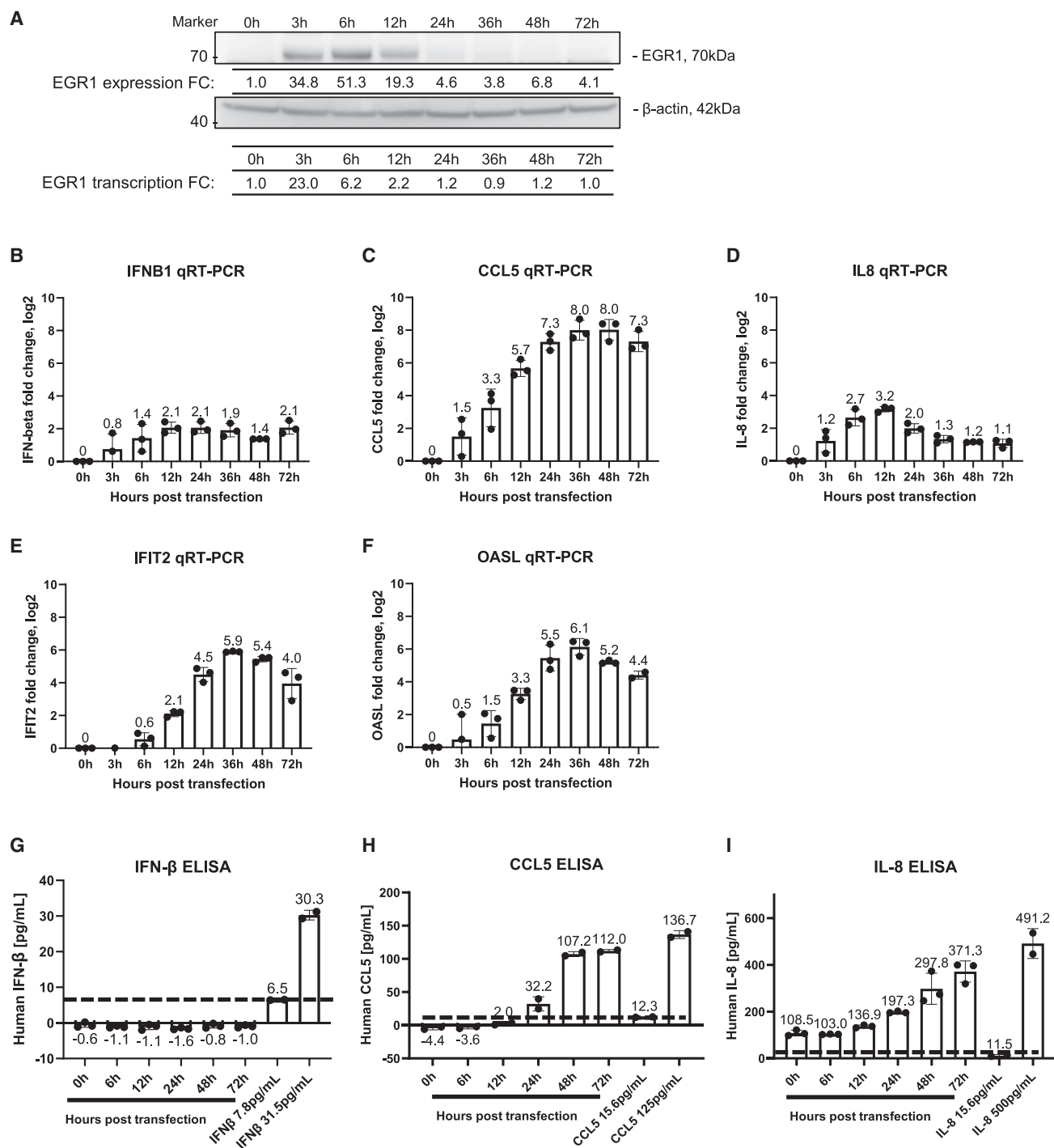
To gain insight into what host pathways may be responsible for the observed pattern of antiviral and inflammatory gene expression, we conducted analysis of potential upstream regulators, on the basis of published gene regulatory databases (Figure 5A). Using this approach on DEGs identified at 72 hpt, 98 potential upstream regulators were predicted to affect 672 DEGs identified at this time point (Figure 5B). This approach was then applied to all seven time points. Note that the upstream regulators are selected regardless of their expression level in the transcriptomic data during the analysis. However, genes that are not detectable prior to or after transfection are unlikely to regulate rAAV production. The union of all seven time points included 107 candidate upstream regulators with expression level higher than 0.5 transcripts per million (TPM) (Figure 5C). The majority of candidate regulators are transcription factors, pathogen recognition receptors, cytokines/cytokine receptors, protein kinases and growth factors. Key signaling nodes of the innate anti-pathogen response were predicted regulators, including interferon alpha/beta receptor subunit 1 (IFNAR1) and interferon regulatory factors (IRFs) IRF1, IRF3, IRF5, and IRF7. Notably, Toll-like receptors, including TLR2 and TLR3, and innate immune signaling transduction adaptor MYD88, as well as both cyclic GMP-AMP synthetase and stimulator of interferon gene, were detectably expressed in transfected HEK293 and predicted as upstream regulators responsible for observed DEGs. This suggests that the process of plasmid transfection and rAAV production directly deliver stimuli that induce anti-pathogen responses in HEK293 cells.

## DISCUSSION

Viral vectors are proving remarkably versatile in multiple approaches to gene therapy, both *in vivo* and *ex vivo*. However, given their distinct requirements for replication and packaging of the therapeutic genetic payload, viral vectors must be made in living cells. Both the

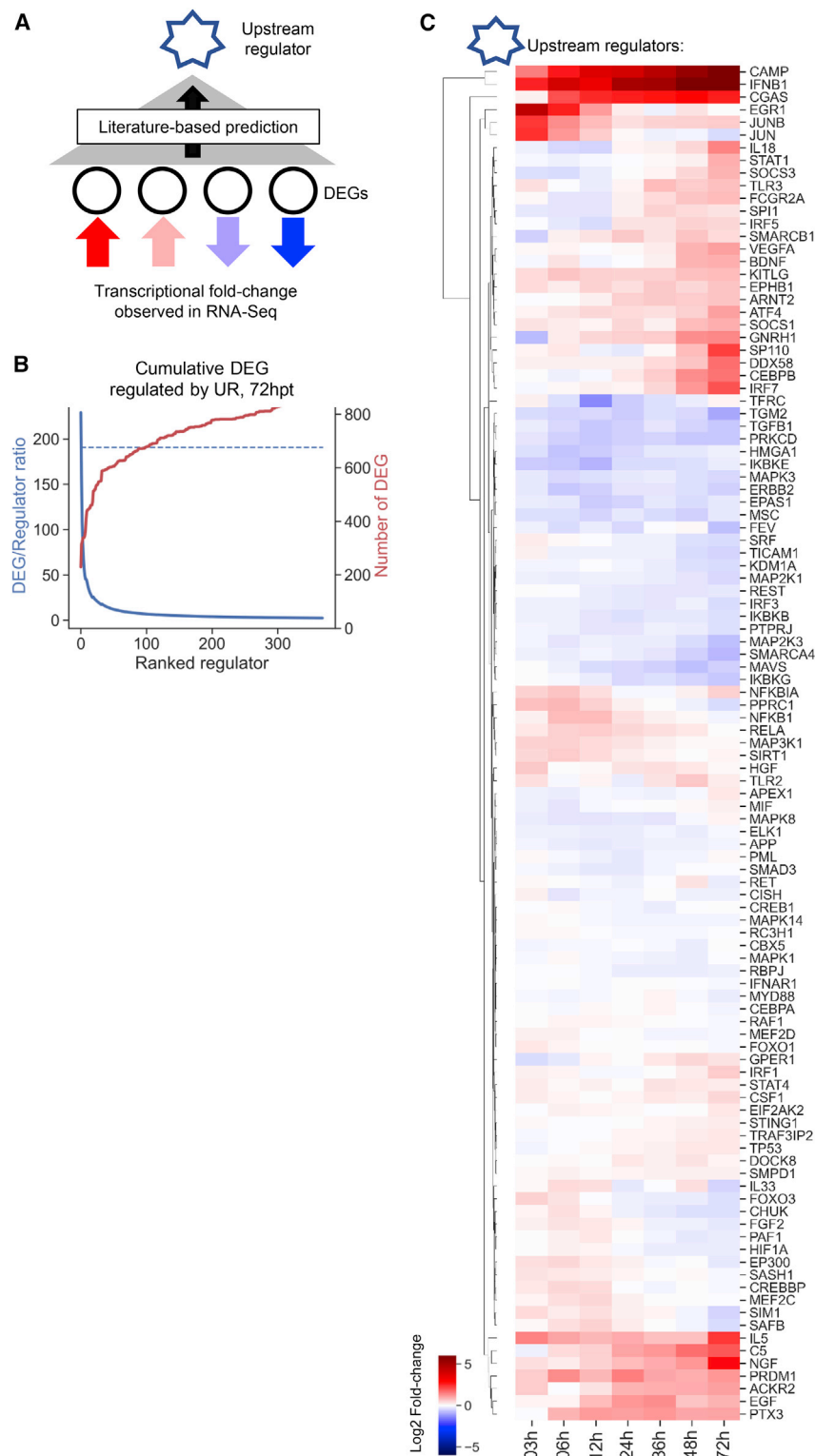
### Figure 3. Overview of transcriptional response to transfection and rAAV production

(A) Number of up- and down-regulated DEGs with at least 2-fold relative to pre-transfection (0 h) baseline. A total of 1,850 DEGs were detected using DESeq2 with FDR < 0.05, as described in materials and methods and Figure S1B. (B and C) Significantly enriched pathways were manually curated and categorized into two major biological functions. The x axis shows the log transformed corrected p value in pathway analysis. GO terms with corrected p values < 0.05 (dotted line) were considered significant, on the basis of Fisher's exact test followed by Benjamini-Hochberg multiple test correction. (D–F) Transcriptional changes at each time point are shown for DEGs associated with Gene Ontology categories for inflammatory response, defense response to virus and nucleosome assembly, respectively. DEGs were clustered by hierarchical clustering on the y axis and ordered by time points on the x axis. Expression fold changes were calculated compared with pre-transfection time point 0 h. The fold changes were indicated in log<sub>2</sub> scale as shown in the lower right legend. (G) DEGs were clustered according to coordinate temporal changes in transcription. The x axis represents sampling time points. The y axis represents log<sub>2</sub> fold change standardized to time zero mean and unit variance across 8 time points and triplicates. Red lines indicate standardized expression level of individual DEGs. The dark blue line represents the mean of clustered genes, while the blue shade area is the range of 95% confidence interval ( $\pm 2$  std. dev.). GO terms were significant when corrected p values were < 0.05 using Fisher's exact test followed by Benjamini-Hochberg multiple test correction.



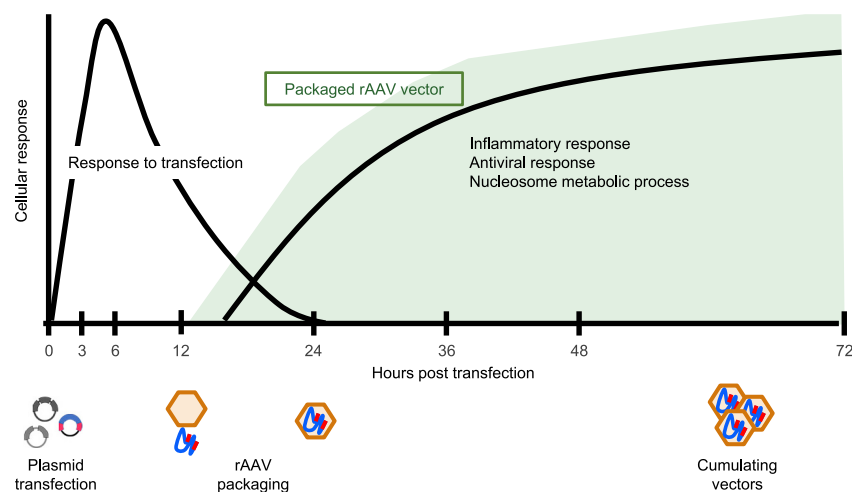
**Figure 4. The up-regulation of selected DEGs are confirmed at protein and mRNA levels**

(A) Time course protein expression of EGR1 in rAAV-producing cells from a 2 L bioreactor by immunoblot. The quantification of EGR1 expression was normalized to actin and compared with pre-transfection (0 h) sample. EGR1 transcription fold change (FC) shows the mean fold change of EGR1 transcripts observed in transcriptome data. (B–F) Select DEGs were confirmed using qRT-PCR analysis of samples from 2 L bioreactors. Time course mRNA expression of IFNβ1 (B), CCL5 (C), IL-8 (D), IFIT2 (E), and OASL (F) are shown. (G–I) Secretion of IFN-β (G), CCL5 (H), and IL-8 (I) in spent medium was quantified using ELISA. The spent medium preparation is described in [materials and methods](#). Lowest linear point of standard curve is shown to indicate the limit of detection (black dashed line). Error bars indicate SD among examined samples.



**Figure 5. Upstream regulator analysis to predict potential master regulators that may modulate transcriptional changes to AAV production**

(A) Schematic of upstream regulator analysis. Using Ingenuity Pathway Analysis (IPA) software, observed expression changes were compared with known gene regulatory interactions, and potential upstream regulators were identified on the basis of consistency between the predicted regulatory effects of each regulator and the actual observed changes in the transcriptome. The circles represent distinct genes. Red arrows (upregulated) and blue arrows (down-regulated) represent the direction of the observed transcriptional changes, while the saturation of arrow colors indicate the magnitude of change on transcription level. (B) The process to select predicted upstream regulators (URs) via IPA analysis was based on the number of modulated DEGs. The predicted URs were ranked by activation Z score, which was calculated in IPA tool inferring whether the expression changes of URs in our RNA-seq agreed (positive Z score) or disagreed (negative Z score) with the protein-protein interaction and/or transcription regulation shown in the literature on the basis of the direction of expression changes of observed DEGs. The UR selection process stopped when the subset of ranked URs collectively regulated more than 80% of DEGs at given time point (dashed line). The red curve indicates the number of cumulative DEGs accounted for at each point in the list of URs. The blue curve represented the DEG/UR ratio given the number of selected URs. (C) The hierarchical clustering heatmap of 107 URs that were expressed in AAV-producing HEK293 cells. Expression fold changes were calculated compared with pre-transfection time point 0 h. The fold changes were indicated in log<sub>2</sub> scale by a hue of red/white/blue color pad that represents upregulated/basal/down-regulated expression level.



**Figure 6. Schematic diagram describing the kinetics of the cellular response to rAAV production in transiently transfected HEK293 cells**

The acute cellular response to transfection is induced at early time points post-transfection, while inflammatory responses and type I IFN signaling pathway are increasingly upregulated over time during rAAV production.

manufacturing processes and the final drug product of viral gene therapy vectors are incompletely understood. Manufacturing highly pure, highly concentrated viral vectors is further complicated by the fact that the mammalian and insect cell lines used to produce them encode multiple pathogen recognition sensors and downstream signaling pathways that evolved to antagonize both exogenous gene expression and viral replication. HEK293 cells have gained broad acceptance as a cellular host for gene therapy vector production, due in part to their ease of transfection. HEK293 express low levels of sensors of exogenous DNA, which may contribute to their utility for viral vector production.<sup>21,33</sup> However, our results indicate that HEK293 are not agnostic to the non-self molecules and parasitic gene expression pathways required to produce viral vectors, but instead respond with robust up-regulation of innate immune response genes (Figure 6). In this study, we identified three cellular responses: (1) negative regulation of MAPK activity, (2) an inflammatory antiviral response, and (3) transcriptional up-regulation of histone genes.

Among these induced cellular responses, antiviral genes clearly have the potential to negatively affect rAAV production.<sup>30</sup> The upregulated DEGs in significantly enriched GOs involving type I interferon and inflammatory response were ISGs inducible not only directly by interferon but also by the interferon regulatory factor family of transcription that are generally activated by pathogen sensors (Figures 3D and 3E).<sup>34</sup> These sensors, also known as pattern recognition receptors, recognize pathogen-associated molecular patterns (PAMPs) from bacterial or viral infection.<sup>35</sup> Analysis of DEGs, combined with literature-based analysis, IRF1/3/5/7 along with PRRs including TLR2, TLR3, cGAS/STING, RIG-I (also known as DDX58), and PKR (also known as EIF2AK2) are implicated as candidate upstream regulators of ISG expression. The key adaptor of TLR signaling, MyD88, was also identified in the list of potential regulators. The PRRs are constitutively expressed for surveillance of exogenous stimuli and are also inducible by antiviral responses. Many TLRs are cell type specific, and most are not generally thought to be expressed in HEK293 cells. However, our data indi-

cated detectable transcription of TLRs 2, 3, 5, and 6. Our analysis suggests that bacterial DNA, viral DNA, or viral RNA in the cytosol, are being recognized by cognate nucleic acid sensors. Although our data do not directly define the pathogen molecules or cellular sensors inducing the gene expression patterns we describe, a recent study demonstrated that

rAAV transduction in mouse induced DNA sensor cGAS and an antiviral response including TNF- $\alpha$ .<sup>36</sup> The potential for endosomal or cytosolic AAV DNA or dsRNA to activate PRRs has not been characterized in rAAV-producing bioreactors. Some HEK293 lines do not express the dominant key adaptor of cytoplasmic DNA sensor, STING, while others are reported to.<sup>22</sup> We found that the HEK293 line used in our experiments do express low but detectable cGAS, which may detect transfected plasmid, or replicating or transduced AAV DNA during bioreactor culture. Multiple additional cytosolic sensors bind transfected plasmid DNA in some cell types, but the relative contribution of these sensors to the effects we observe in HEK293 is not known.<sup>37</sup> TLR2, another potential regulator identified by our analyses, may be activated by bacterial cell wall components such as triacylated lipoproteins.<sup>38</sup> The ligands for TLR2 activation could come from the residual bacterial lipoproteins derived from the *Escherichia coli* plasmid host. Finally, AAV capsid has been reported as a PAMP recognized by TLR2 on the plasma membrane and induced pro-inflammatory cytokines through NF- $\kappa$ B signaling.<sup>31</sup> Note that the sensors and adaptors hypothesized above were transcriptionally expressed on the basis of our RNA-seq data. This contradicts previous evidence that TLR family, cGAS, STING, or RIG-I were not detectable in HEK293 cells.<sup>21-23</sup> Consistent with our findings, a recent study identified “defense response” as one of the significantly enriched pathways in a proteomic analysis comparing transfected AAV5-producing HEK293 cells to non-transfected controls, likely induced by stress including transfection or exogenous components.<sup>39</sup> Additional investigation is required to elucidate exactly which sensors cause the induction of the observed antiviral response in HEK293 cells during rAAV production.

The MAPK family is one of the earliest kinase cascades activated in response to various stressors, including pathogen infection. Previous studies suggested that the internalization of polyethylenimine transfection reagent could induce cell stress and trigger PKC activation.<sup>13</sup> The stress response activates MAPK signaling pathway and



transient up-regulation of immediate-early genes including EGR1 and FOS/JUN gene families, as well as DUSPs that properly control the MAPK activation.<sup>27,28</sup> The up-regulation of the DUSPs is a negative feedback mechanism to prevent sustained MAPK activation, and several DUSPs have been implicated in down-regulating the signaling cascades activated by PRR-PAMP interactions.<sup>28</sup> Previous research that directly compared transfected and non-transfected cells demonstrated that transient transfection of plasmid DNA activated DNA sensors and induced an innate immune response.<sup>29,40-42</sup> Regulation of MAPK activation and up-regulation of MAPK inducible genes such as JUN were observed after transient transfection of plasmids for protein production of non-viral products, aligning with our hypothesis of MAPK activation due to transient transfection. Notably, in prior studies, the gene expression profiles following transfection of plasmid are highly dependent on cell type, transfection reagent, culture process conditions, and the nature of the exogenous proteins transcribed.<sup>40-42</sup> This observation underscores the need to perform transcriptomic and proteomic studies under manufacturing-relevant conditions. Regardless, the early cellular response following transfection is a potential target for bioengineering to improve viral vector manufacturing.

We observed up-regulation of 39 histone genes with mid-late kinetics, at 24 hpt. Histone gene transcription is significantly increased during S phase in mammalian cells.<sup>43</sup> Histone gene transcription is tightly coupled with the progression of DNA synthesis and down-regulated as cells exit S phase or by replication inhibitors.<sup>44,45</sup> Therefore, histone up-regulation is a characteristic of high energy demand tied with active cell proliferation or, in the case of rAAV production, likely with intensive vector genome replication. Previous proteomic analysis on transfected HEK293 cells producing virus-like particles identified significant up-regulation of histones and mitochondrial components under conditions of high metabolic demand, which is consistent with our findings of increased histone expression.<sup>46</sup> The adenovirus genes E1A and E4orf6 that are required for effective rAAV production inhibit Cdc2 and promote degradation of cyclin A, allow HEK293 cells stay in S phase.<sup>47</sup> Therefore, the histone up-regulation could be a consequence of cell cycle regulation induced by helper proteins that facilitate AAV genome replication or vector production.

Production of biotherapeutics is heavily dependent upon the choice of producer cell line for efficient manufacture. In the case of monoclonal antibody (mAb) production, highly efficient CHO cell lines have been developed by adaptation and targeted modification.<sup>48</sup> In contrast to mAbs and other recombinant protein products, the interaction between the host cell and viral replication during the manufacture of viral vectors is significantly more complicated. The AAV Rep proteins and helper genes E2a and E4 are cytotoxic.<sup>49-51</sup> In addition to producing the protein components of the vector, the cells must also host a robust viral DNA replication process. Viral replication frequently triggers an inhibitory innate immune response, which we demonstrate here is the case for rAAV production following plasmid transfection. The antiviral response has been shown to affect protein synthesis, cellular metabolism, cell proliferation and even promote

programmed cell death, and these responses could significantly curtail rAAV production in HEK293 cells.<sup>52</sup> Overexpression of cGAS/STING to activate DNA sensing response in lentiviral vector-producing HEK293T cells activated promoters of NF- $\kappa$ B and interferon- $\beta$  promoters.<sup>53</sup> The activation of the downstream innate immune response, however, did not strongly affect lentiviral vector production. Nevertheless, existing pathogen sensors that are constitutively expressed may play more significant roles in the regulation of vector production in these producer cell lines. A recent study improved the production of lentiviral vector in HEK293 cells by 6.7-fold after knocking out constitutively expressed antiviral effectors including PKR and 2'-5'-oligoadenylate synthetase 1 (OAS1), suggesting that host pathways were capable of inhibiting vector production.<sup>54</sup> This is consistent with our finding on the induction of an antiviral response during rAAV production. Together, the time course transcriptomic analysis revealed the cellular response and potential pathways that the host cells use to recognize exogenous stimuli throughout the process of rAAV vector production in suspension HEK293 cells at manufacturing scale for the first time. Improved understanding of the state of the host cell during rAAV production may lead to improvement of vector production platforms that are better able to meet the growing clinical demand for these therapeutics.

## MATERIALS AND METHODS

### HEK293 cell line and cell culture

For the cell expansion for 50 L bioreactor culture, an HEK293-derived suspension cell line was cultured in 1 L serum-free FreeStyle F17 Expression Medium supplemented with GlutaMAX Supplement at 10 mM (Thermo Fisher Scientific) at 37°C, 5% CO<sub>2</sub>, 50% humidity, and 64 rpm in a 3 L Corning vent cap flask. Cells were then seeded at 0.3E+6 viable cell density in 5 L of F17 medium supplemented with GlutaMAX at 10 mM, 0.2% pluronic F68 and 0.1% FoamAway Irradiated Animal Origin-Free antifoaming agent (Thermo Fisher Scientific) in a 10 L Wave bag. The Wave bag was rocked at 25 rpm, 37°C, 5% CO<sub>2</sub>. Finally, cells were cultured in 50 L single-use bioreactor (Thermo Fisher Scientific) containing 20 L of F17 supplemented with 10 mM GlutaMAX, 500 mL of 10% pluronic F68 and 25 mL of FoamAway and 5 L cell culture medium from 10 L Wave bag described above. The temperature was set at 37°C, and the agitator was set at 70 rpm. Forty-eight hours after inoculation, 15 L of Expi293 Expression medium (Thermo Fisher Scientific) along with 300 mL of 10% pluronic F68 were added to the culture. The culture was brought to cell density of 5E+6 cells/mL before transfection. Viable cell density (VCD) and cell viability were measured using a Vi-Cell XR cell viability analyzer (Beckman Coulter Life Sciences). For the culture in 2 L bioreactors, the medium and process controls were remained the same as 50 L SUBs and targeted the final culture volume of 1 L with cell density of 5E+6 cells/mL before transfection.

### Transfection for rAAV production

Plasmid mix including helper pXX680, Rep/Cap pGSK2/9 and transgene pAAV-OptiDys at 1:2:1 molar ratio was transfected using polyethylenimine transfection reagent. pAAV-OptiDys used here is

a transgene plasmid that expresses a human mini-dystrophin gene, transcribed by a skeletal muscle-specific modified human creatine kinase promoter and enhancer, and a synthetic poly adenylation signal. The transgene is described more fully in published international patent application WO2017221145. The length of packaged AAV genome is 4.95 kb, including ITRs. One microgram of plasmid mix per million viable cells was used for cell transfection. PEI was added at two times the mass of the three-plasmid mix, seven minutes before delivering the transfection mix to the SUB or 2 L bioreactor. Viable cell density and percentage of viable cells were measured using Vi-Cell at the 8 selected sampling time points.

#### rAAV titer determination

For 50 L SUB samples, cells sampled at 36, 48, and 72 h after transfection were stored at  $-80^{\circ}\text{C}$  and then sonicated for 4 min with a one second pulse between every second of sonication at 20% power (Branson Digital). For 2 L bioreactor samples, lysis was achieved by three freeze-thaw cycles. Unencapsidated DNA in cell lysate was digested using DNase-I (MP Biomedicals) at 263 U/mL incubated at room temperature for 60 min, followed by DNase-I inactivation and rAAV capsid proteolytic degradation at the same time using 1.3 mM EDTA and 0.53X of Teknova qPCR protease K buffer (1.11 M NaCl, 1.11% Sarkosyl) at  $95^{\circ}\text{C}$  incubation for 10 min. FAM-labeled probe and primer pair specific for ITR region (for samples from 50 L SUBs) or OptiDys transgene ORF region (for samples from 2 L bioreactors) were used for TaqMan qPCR protocol.

#### RNA-seq

Cell pellets at 8 selected sampling time points were submitted for RNA-seq. mRNA enrichment followed by random priming was used for RNA library preparation to exclude rRNA or tRNA. Each RNA library was sequenced three times by Illumina HiSeq  $2 \times 150$  pair-end configuration. The RNA library preparation and sequencing were conducted by GENEWIZ Inc. Note that each sample was submitted to individual library preparation and independent flow cells without multiplexing, eliminating the likelihood that the clustering was due to artifact of next-generation sequencing (NGS).

#### Transcriptome data preprocessing

Expression levels of 35,616 distinct transcripts in the human genome were determined by Kallisto using RNA-seq data.<sup>55</sup> The reference genome sequences for read mapping was GRCh38.99 along with viral transcripts on either HEK293 cells including E1A and E1B or three transfected plasmids including E2A, E4orf6, VAI, VAI1, p5 transcript, p19 transcript, and p40 transcript. Three technical repeats from each RNA library were pooled to determine expression level for each transcript. Transcripts with more than 0.5 transcripts per million at any time points were defined as expressed transcripts. The TPM cutoff was set to observe at least one transcript in each sampled cell assuming two million RNA molecules per cell in average at the time of sampling.

#### Bioinformatic analysis

Differentially expressed genes were identified using likelihood ratio test (LRT) in DESeq2 package in R for any expressed transcript

that show significant change in expression across 8 time points with false discovery rate-adjusted p value less than 0.05 using custom R script. Time-series analysis was performed using Python program based on a Dirichlet process Gaussian process mixture model for expression pattern clustering.<sup>25</sup> The DPGP mixture model assumes that the expression levels at adjacent time points are dependent, which makes an ideal fit for time course transcriptome samples. DEGs were submitted to clustering with 200 iterations of cluster assignment using DPGP mixture model. Enriched pathway analysis was conducted using GOATOOLS.<sup>56</sup> The results and figures generated from *DP\_GP* and *goatools* Python packages can be reproduced using custom Python code.

#### qRT-PCR

Total RNA was extracted using TRIzol reagent (Invitrogen). Reverse transcription was achieved with oligo dT primer using SuperScript IV first-strand synthesis system (Invitrogen) for enrichment of polyA<sup>+</sup> messenger RNAs. Expression level of target genes was measured by qPCR using FAM-labeled TaqMan Gene Expression Assay (IFNB1: Hs01077958\_s1; CCL5: Hs00982282\_m1; CXCL8: Hs00174103\_m1; OASL: Hs00984387\_m1; IFIT2: Hs00533665\_m1) and TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific). HPRT1 (Hs02800695\_m1) was used as internal control.

#### Enzyme-linked immunosorbent assay

The time course cell culture samples post-transfection were collected from three independent bioreactors. The spent medium was collected from the supernatant of culture after centrifugation at  $2,000 \times g$  for 10 min at  $4^{\circ}\text{C}$ . Spent medium was concentrated 10-fold using disposable ultrafiltration units with a molecular weight cutoff of 3 kDa (Thermo Fisher Scientific) before assay. Standard curve was diluted in concentrated medium (F17 and Expi293 in 1:1 ratio) to account for potential matrix effect of concentrated medium selected ELISA kits. The limit of detection was set at the point with lowest concentration on the linear slope of the standard curve and the absorbance obtained using concentrated medium negative control was defined as noise for background subtraction.

#### DATA AVAILABILITY

The raw sequencing data will be deposited at Gene Expression Omnibus.

#### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtm.2023.01.002>.

#### ACKNOWLEDGMENTS

This project is supported by the Pfizer Worldwide Research, Development, and Medical Postdoctoral training program.

#### AUTHOR CONTRIBUTIONS

Experimental design or interpretation of results: E.B., C.-H.C., D.M., C.M.M., J.W.P., and J.R.; technical assistance: R.N., W.W., and

V.P.W.; prepared manuscript: C.-H.C., C.M.M., and E.B.; critical reading and revision: D.M., R.N., V.P.W., J.P., and E.B.

## DECLARATION OF INTERESTS

All authors were employed by Pfizer Inc. when the present research was conducted.

## REFERENCES

- Penaud-Budloo, M., François, A., Clément, N., and Ayuso, E. (2018). Pharmacology of recombinant adeno-associated virus production. *Mol. Ther. Methods Clin. Dev.* 8, 166–180. <https://doi.org/10.1016/j.omtm.2018.01.002>.
- Xiao, X., Li, J., and Samulski, R.J. (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J. Virol.* 72, 2224–2232. <https://doi.org/10.1128/JVI.72.3.2224-2232.1998>.
- Berk, A.J. (2005). Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. *Oncogene* 24, 7673–7685. <https://doi.org/10.1038/sj.onc.1209040>.
- Louis, N., Eveleigh, C., and Graham, F.L. (1997). Cloning and sequencing of the cellular-viral junctions from the human adenovirus type 5 transformed 293 cell line. *Virology* 233, 423–429. <https://doi.org/10.1006/viro.1997.8597>.
- Graham, F.L., Smiley, J., Russell, W.C., and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36, 59–74. <https://doi.org/10.1099/0022-1317-36-1-59>.
- Côté, J., Garnier, A., Massie, B., and Kamen, A. (1998). Serum-free production of recombinant proteins and adenoviral vectors by 293SF-3F6 cells. *Biotechnol. Bioeng.* 59, 567–575.
- Garnier, A., Côté, J., Nadeau, I., Kamen, A., and Massie, B. (1994). Scale-up of the adenovirus expression system for the production of recombinant protein in human 293S cells. *Cytotechnology* 15, 145–155. <https://doi.org/10.1007/BF00762389>.
- Graham, F.L. (1987). Growth of 293 cells in suspension culture. *J. Gen. Virol.* 68, 937–940. <https://doi.org/10.1099/0022-1317-68-3-937>.
- Grieger, J.C., Soltys, S.M., and Samulski, R.J. (2016). Production of recombinant adeno-associated virus vectors using suspension HEK293 cells and continuous harvest of vector from the culture media for GMP FIX and FLT1 clinical vector. *Mol. Ther.* 24, 287–297. <https://doi.org/10.1038/mt.2015.187>.
- FDA, U.S. (2021). Cellular, tissue, and gene therapies advisory committee september 2-3, 2021 meeting briefing document- FDA. In CBER.
- Raof, N.A., Rajamani, D., Chu, H.C., Gurav, A., Johnson, J.M., LoGerfo, F.W., Pradhan-Nabzdyk, L., and Bhasin, M. (2016). The effects of transfection reagent polyethyleneimine (PEI) and non-targeting control siRNAs on global gene expression in human aortic smooth muscle cells. *BMC Genom.* 17, 20. <https://doi.org/10.1186/s12864-015-2267-9>.
- Brunot, C., Ponsonnet, L., Lagneau, C., Farge, P., Picart, C., and Grosogeat, B. (2007). Cytotoxicity of polyethyleneimine (PEI), precursor base layer of polyelectrolyte multilayer films. *Biomaterials* 28, 632–640. <https://doi.org/10.1016/j.biomaterials.2006.09.026>.
- Kopatz, I., Remy, J.S., and Behr, J.P. (2004). A model for non-viral gene delivery: through syndecan adhesion molecules and powered by actin. *J. Gene Med.* 6, 769–776. <https://doi.org/10.1002/jgm.558>.
- Berthet, C., Raj, K., Saudan, P., and Beard, P. (2005). How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proc. Natl. Acad. Sci. USA* 102, 13634–13639. <https://doi.org/10.1073/pnas.0504583102>.
- Hermanns, J., Schulze, A., Jansen-Db1urr, P., Kleinschmidt, J.A., Schmidt, R., and zur Hausen, H. (1997). Infection of primary cells by adeno-associated virus type 2 results in a modulation of cell cycle-regulating proteins. *J. Virol.* 71, 6020–6027. <https://doi.org/10.1128/JVI.71.8.6020-6027.1997>.
- Zhu, J., Huang, X., and Yang, Y. (2009). The TLR9-MyD88 pathway is critical for adaptive immune responses to adeno-associated virus gene therapy vectors in mice. *J. Clin. Invest.* 119, 2388–2398. <https://doi.org/10.1172/JCI37607>.
- Martino, A.T., Suzuki, M., Markusic, D.M., Zolotukhin, I., Ryals, R.C., Moghimi, B., Ertl, H.C.J., Muruve, D.A., Lee, B., and Herzog, R.W. (2011). The genome of self-complementary adeno-associated viral vectors increases Toll-like receptor 9-dependent innate immune responses in the liver. *Blood* 117, 6459–6468. <https://doi.org/10.1182/blood-2010-10-314518>.
- Faust, S.M., Bell, P., Cutler, B.J., Ashley, S.N., Zhu, Y., Rabinowitz, J.E., and Wilson, J.M. (2013). CpG-depleted adeno-associated virus vectors evade immune detection. *J. Clin. Invest.* 123, 2994–3001. <https://doi.org/10.1172/JCI68205>.
- Stracker, T.H., Carson, C.T., and Weitzman, M.D. (2002). Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex. *Nature* 418, 348–352. <https://doi.org/10.1038/nature00863>.
- Hart, L.S., Ornelles, D., and Koumenis, C. (2007). The adenoviral E4orf6 protein induces atypical apoptosis in response to DNA damage. *J. Biol. Chem.* 282, 6061–6067. <https://doi.org/10.1074/jbc.M610405200>.
- Sun, L., Wu, J., Du, F., Chen, X., and Chen, Z.J. (2013). Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* 339, 786–791. <https://doi.org/10.1126/science.1232458>.
- Burdette, D.L., Monroe, K.M., Sotelo-Troha, K., Iwig, J.S., Eckert, B., Hyodo, M., Hayakawa, Y., and Vance, R.E. (2011). STING is a direct innate immune sensor of cyclic di-GMP. *Nature* 478, 515–518. <https://doi.org/10.1038/nature10429>.
- Hornung, V., Rothenfusser, S., Britsch, S., Krug, A., Jahrsdörfer, B., Giese, T., Endres, S., and Hartmann, G. (2002). Quantitative expression of toll-like receptor 1-10 mRNA in cellular subsets of human peripheral blood mononuclear cells and sensitivity to CpG oligodeoxynucleotides. *J. Immunol.* 168, 4531–4537. <https://doi.org/10.4049/jimmunol.168.9.4531>.
- Catterall, T., Fynch, S., Kay, T.W.H., Thomas, H.E., and Sutherland, A.P.R. (2020). IL-17F induces inflammation, dysfunction and cell death in mouse islets. *Sci. Rep.* 10, 13077. <https://doi.org/10.1038/s41598-020-69805-2>.
- McDowell, I.C., Manandhar, D., Vockley, C.M., Schmid, A.K., Reddy, T.E., and Engelhardt, B.E. (2018). Clustering gene expression time series data using an infinite Gaussian process mixture model. *PLoS Comput. Biol.* 14, e1005896. <https://doi.org/10.1371/journal.pcbi.1005896>.
- Rivals, I., Personnaz, L., Taing, L., and Potier, M.C. (2007). Enrichment or depletion of a GO category within a class of genes: which test? *Bioinformatics* 23, 401–407. <https://doi.org/10.1093/bioinformatics/btl633>.
- Bahrani, S., and Drablos, F. (2016). Gene regulation in the immediate-early response process. *Adv. Biol. Regul.* 62, 37–49. <https://doi.org/10.1016/j.jbior.2016.05.001>.
- Lang, R., and Raffi, F.A.M. (2019). Dual-specificity phosphatases in immunity and infection: an update. *Int. J. Mol. Sci.* 20, 2710. <https://doi.org/10.3390/ijms20112710>.
- Li, S., Wilkinson, M., Xia, X., David, M., Xu, L., Purkel-Sutton, A., and Bhardwaj, A. (2005). Induction of IFN-regulated factors and antitumoral surveillance by transfected placebo plasmid DNA. *Mol. Ther.* 11, 112–119. <https://doi.org/10.1016/j.ymthe.2004.09.008>.
- Liu, J., Guan, X., and Ma, X. (2005). Interferon regulatory factor 1 is an essential and direct transcriptional activator for interferon  $\gamma$ -induced RANTES/CCL5 expression in macrophages. *J. Biol. Chem.* 280, 24347–24355. <https://doi.org/10.1074/jbc.M500973200>.
- Lee, A.H., Hong, J.H., and Seo, Y.S. (2000). Tumour necrosis factor- $\alpha$  and interferon- $\gamma$  synergistically activate the RANTES promoter through nuclear factor  $\kappa$ B and interferon regulatory factor 1 (IRF-1) transcription factors. *Biochem. J.* 350, 131–138.
- Wagoner, J., Austin, M., Green, J., Imaizumi, T., Casola, A., Brasier, A., Khabar, K.S.A., Wakita, T., Gale, M., Jr., and Polyak, S.J. (2007). Regulation of CXCL-8 (interleukin-8) induction by double-stranded RNA signaling pathways during hepatitis C virus infection. *J. Virol.* 81, 309–318. <https://doi.org/10.1128/JVI.01411-06>.
- Igoucheva, O., Alexeev, V., and Yoon, K. (2006). Differential cellular responses to exogenous DNA in mammalian cells and its effect on oligonucleotide-directed gene modification. *Gene Ther.* 13, 266–275. <https://doi.org/10.1038/sj.gt.3302643>.
- Taniguchi, T., Ogasawara, K., Takaoka, A., and Tanaka, N. (2001). IRF family of transcription factors as regulators of host defense. *Annu. Rev. Immunol.* 19, 623–655. <https://doi.org/10.1146/annurev.immunol.19.1.623>.
- Gordon, S. (2002). Pattern recognition receptors: doubling up for the innate immune response. *Cell* 111, 927–930. [https://doi.org/10.1016/s0092-8674\(02\)01201-1](https://doi.org/10.1016/s0092-8674(02)01201-1).
- Chandler, L.C., Barnard, A.R., Caddy, S.L., Patrício, M.I., McClements, M.E., Fu, H., Rada, C., MacLaren, R.E., and Xue, K. (2019). Enhancement of adeno-associated

- virus-mediated gene therapy using hydroxychloroquine in murine and human tissues. *Mol. Ther. Methods Clin. Dev.* 14, 77–89. <https://doi.org/10.1016/j.omtm.2019.05.012>.
37. Semenova, N., Bosnjak, M., Markelc, B., Znidar, K., Cemazar, M., and Heller, L. (2019). Multiple cytosolic DNA sensors bind plasmid DNA after transfection. *Nucleic Acids Res.* 47, 10235–10246. <https://doi.org/10.1093/nar/gkz768>.
38. Aliprantis, A.O., Yang, R.B., Mark, M.R., Suggett, S., Devaux, B., Radolf, J.D., Klimpel, G.R., Godowski, P., and Zychlinsky, A. (1999). Cell activation and apoptosis by bacterial lipoproteins through toll-like receptor-2. *Science* 285, 736–739. <https://doi.org/10.1126/science.285.5428.736>.
39. Strasser, L., Boi, S., Guapo, F., Donohue, N., Barron, N., Rainbow-Fletcher, A., and Bones, J. (2021). Proteomic landscape of adeno-associated virus (AAV)-Producing HEK293 cells. *Int. J. Mol. Sci.* 22, 11499. <https://doi.org/10.3390/ijms222111499>.
40. Jacobsen, L., Calvin, S., and Lobenhofer, E. (2009). Transcriptional effects of transfection: the potential for misinterpretation of gene expression data generated from transiently transfected cells. *Biotechniques* 47, 617–624. <https://doi.org/10.2144/000113132>.
41. Waga, E., Anderson, J., Tucker, M., Harris, E., and Elmer, J. (2022). Transcriptomic analysis of the innate immune response to in vitro transfection of plasmid DNA. *Molecular Therapy - Nucleic Acids*. <https://doi.org/10.1016/j.omtn.2022.11.025>.
42. Stepanenko, A.A., and Heng, H.H. (2017). Transient and stable vector transfection: pitfalls, off-target effects, artifacts. *Mutat. Res. Rev. Mutat. Res.* 773, 91–103. <https://doi.org/10.1016/j.mrrev.2017.05.002>.
43. DeLisle, A.J., Graves, R.A., Marzluff, W.F., and Johnson, L.F. (1983). Regulation of histone mRNA production and stability in serum-stimulated mouse 3T6 fibroblasts. *Mol. Cell Biol.* 3, 1920–1929. <https://doi.org/10.1128/mcb.3.11.1920-1929.1983>.
44. Polo, S.E., and Almouzni, G. (2005). Histone metabolic pathways and chromatin assembly factors as proliferation markers. *Cancer Lett.* 220, 1–9. <https://doi.org/10.1016/j.canlet.2004.08.024>.
45. Marzluff, W.F., and Duronio, R.J. (2002). Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. *Curr. Opin. Cell Biol.* 14, 692–699. [https://doi.org/10.1016/s0955-0674\(02\)00387-3](https://doi.org/10.1016/s0955-0674(02)00387-3).
46. Lavado-García, J., Jorge, I., Cervera, L., Vázquez, J., and Gódia, F. (2020). Multiplexed quantitative proteomic analysis of HEK293 provides insights into molecular changes associated with the cell density effect, transient transfection, and virus-like particle production. *J. Proteome Res.* 19, 1085–1099. <https://doi.org/10.1021/acs.jproteome.9b00601>.
47. Ben-Israel, H., and Kleinberger, T. (2002). Adenovirus and cell cycle control. *Front. Biosci.* 7, d1369–d1395. <https://doi.org/10.2741/ben>.
48. Fischer, S., Handrick, R., and Otte, K. (2015). The art of CHO cell engineering: a comprehensive retrospect and future perspectives. *Biotechnol. Adv.* 33, 1878–1896. <https://doi.org/10.1016/j.biotechadv.2015.10.015>.
49. Schmidt, M., Afione, S., and Kotin, R.M. (2000). Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53. *J. Virol.* 74, 9441–9450. <https://doi.org/10.1128/jvi.74.20.9441-9450.2000>.
50. Jing, X.J., Kalman-Maltese, V., Cao, X., Yang, Q., and Trempe, J.P. (2001). Inhibition of adenovirus cytotoxicity, replication, and E2a gene expression by adeno-associated virus. *Virology* 291, 140–151. <https://doi.org/10.1006/viro.2001.1192>.
51. Marcellus, R.C., Lavoie, J.N., Boivin, D., Shore, G.C., Ketner, G., and Branton, P.E. (1998). The early region 4 orf4 protein of human adenovirus type 5 induces p53-independent cell death by apoptosis. *J. Virol.* 72, 7144–7153. <https://doi.org/10.1128/JVI.72.9.7144-7153.1998>.
52. Fritsch, S.D., and Weichhart, T. (2016). Effects of interferons and viruses on metabolism. *Front. Immunol.* 7, 630. <https://doi.org/10.3389/fimmu.2016.00630>.
53. Ferreira, C.B., Sumner, R.P., Rodríguez-Plata, M.T., Rasaiyaah, J., Milne, R.S., Thrasher, A.J., Qasim, W., and Towers, G.J. (2020). Lentiviral vector production titer is not limited in HEK293T by induced intracellular innate immunity. *Mol. Ther. Methods Clin. Dev.* 17, 209–219. <https://doi.org/10.1016/j.omtm.2019.11.021>.
54. Han, J., Tam, K., Tam, C., Hollis, R.P., and Kohn, D.B. (2021). Improved lentiviral vector titers from a multi-gene knockout packaging line. *Mol. Ther. Oncolytics* 23, 582–592. <https://doi.org/10.1016/j.omto.2021.11.012>.
55. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527. <https://doi.org/10.1038/nbt.3519>.
56. Klopfenstein, D.V., Zhang, L., Pedersen, B.S., Ramírez, F., Warwick Vesztrocy, A., Naldi, A., Mungall, C.J., Yunes, J.M., Botvinnik, O., Weigel, M., et al. (2018). Goatools: a Python library for Gene Ontology analyses. *Sci. Rep.* 8, 10872. <https://doi.org/10.1038/s41598-018-28948-z>.

**Supplemental information**

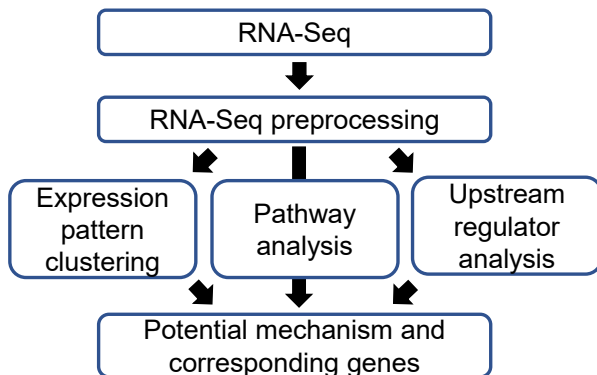
**Production of rAAV by plasmid transfection  
induces antiviral and inflammatory  
responses in suspension HEK293 cells**

**Cheng-Han Chung, Christopher M. Murphy, Vincent P. Wingate, Jeffrey W. Pavlicek, Reiko Nakashima, Wei Wei, Douglas McCarty, Joseph Rabinowitz, and Erik Barton**



**A**

Bioinformatic workflow:

**B**

RNA-Seq preprocessing:

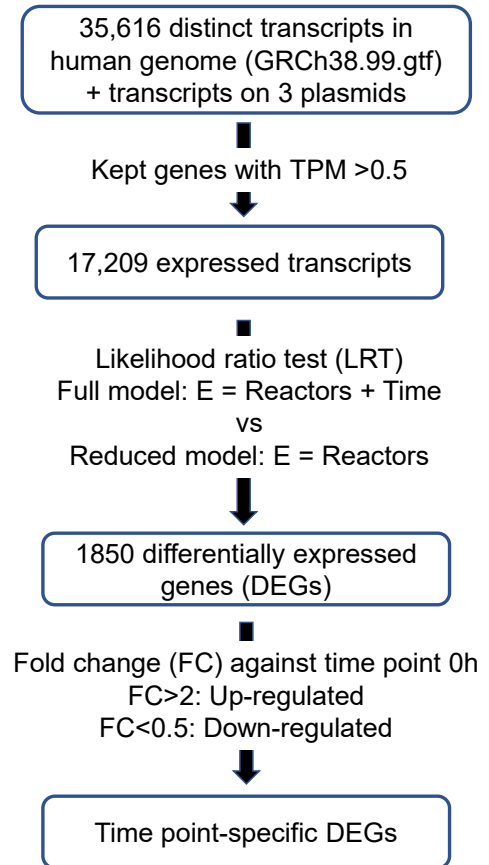


Figure S1. (A) Bioinformatic workflow for the identification of potential mechanisms and corresponding genes that regulated AAV production via plasmid transfection in HEK293 cells. A total of 708,113 Mbases was yield across 24 samples with mean quality score of 35.85. The percentage of sequences with error rate lower than 0.001 was 93.17%. Expression pattern clustering was performed using time-series analysis based on Dirichlet process Gaussian process (DPGP) mixture model <sup>1</sup>. Pathway analysis was conducted using GOATOOLS with Gene Ontology (GO) database <sup>2</sup>. The upstream regulator analysis was one function in Ingenuity Pathway Analysis (IPA) that identifies upstream transcriptional regulators that coordinate the expressional changes observed in transcriptome. (B) Step-by-step RNA-Seq preprocessing conducted in this study and corresponding criteria to select regulated genes by transcriptional changes. TPM: transcript per million, estimated by kallisto <sup>3</sup>.

## PCA plot using 17,209 expressed transcripts

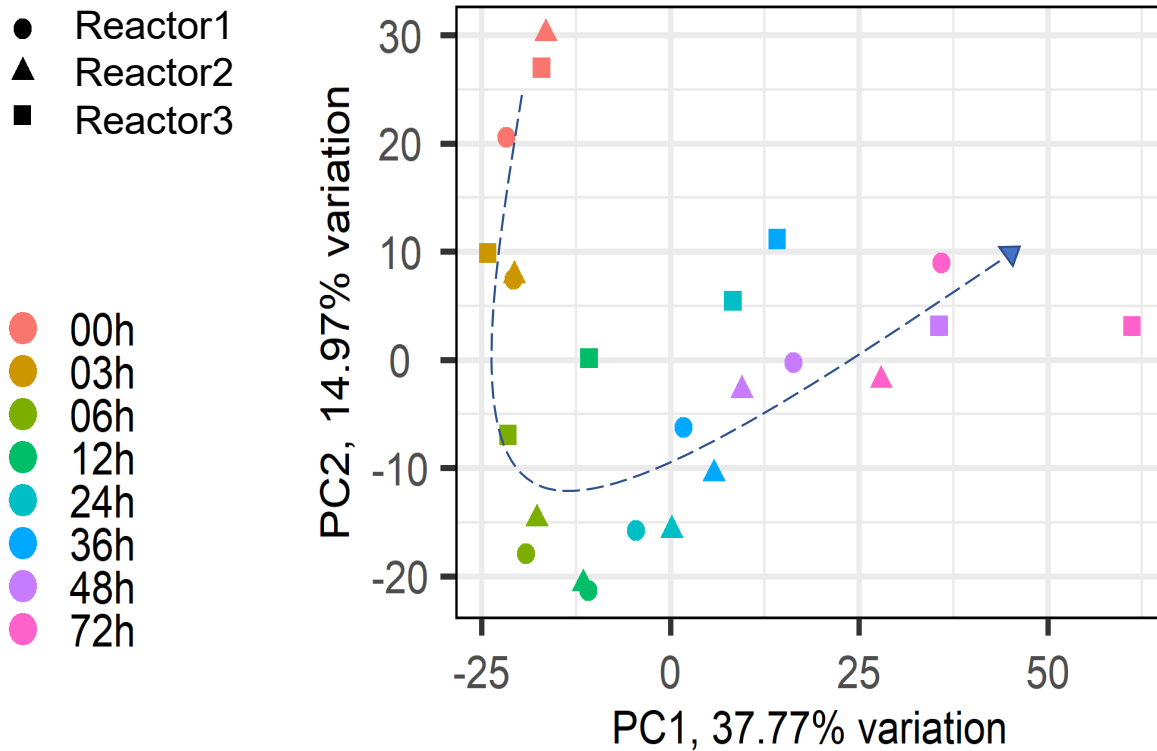


Figure S2. Transcriptome data showed high reproducibility among replicates within sampling time points and progressing expression profiles in relation to time in principal component analysis (PCA). The dotted line indicates the trendline of transcriptional changes over time on the plane of PC1 and PC2. These two principal components combined can explain 52.8% of variation of transcriptional change in RNA-Seq results.

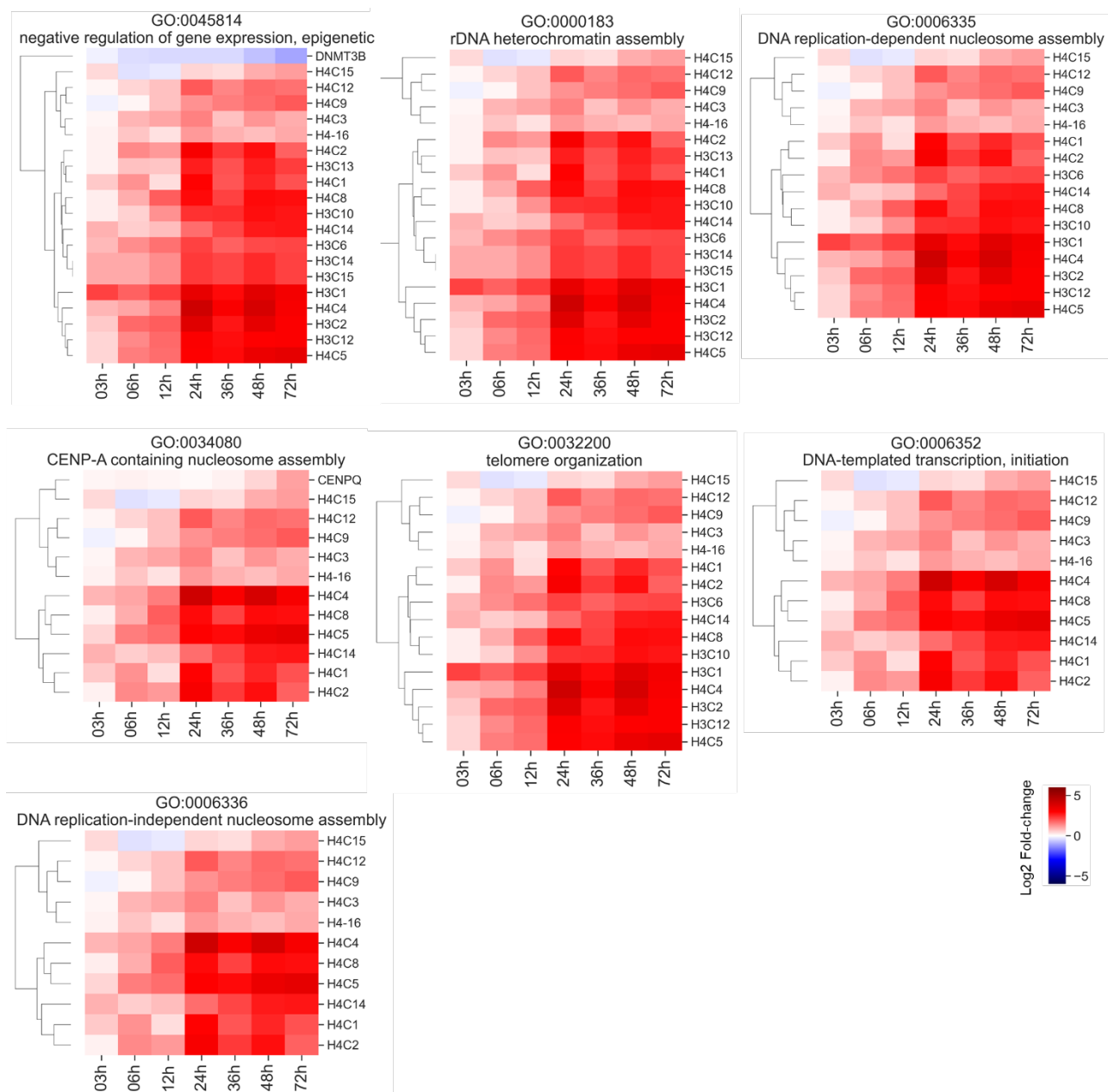


Figure S3. Hierarchical clustering of genes involved in nucleosome assembly and 8 related biological processes significantly enriched in GO enrichment analysis. Heatmap was ordered by time points. All listed genes were DEGs that passed the selection criteria described in Figure S1B. Expression fold-changes were calculated compared with pre-transfection time point 0h. The fold changes were indicated in log2 scale by a hue of red/white/blue color pad that represents up-regulated/basal/down-regulated expression level.

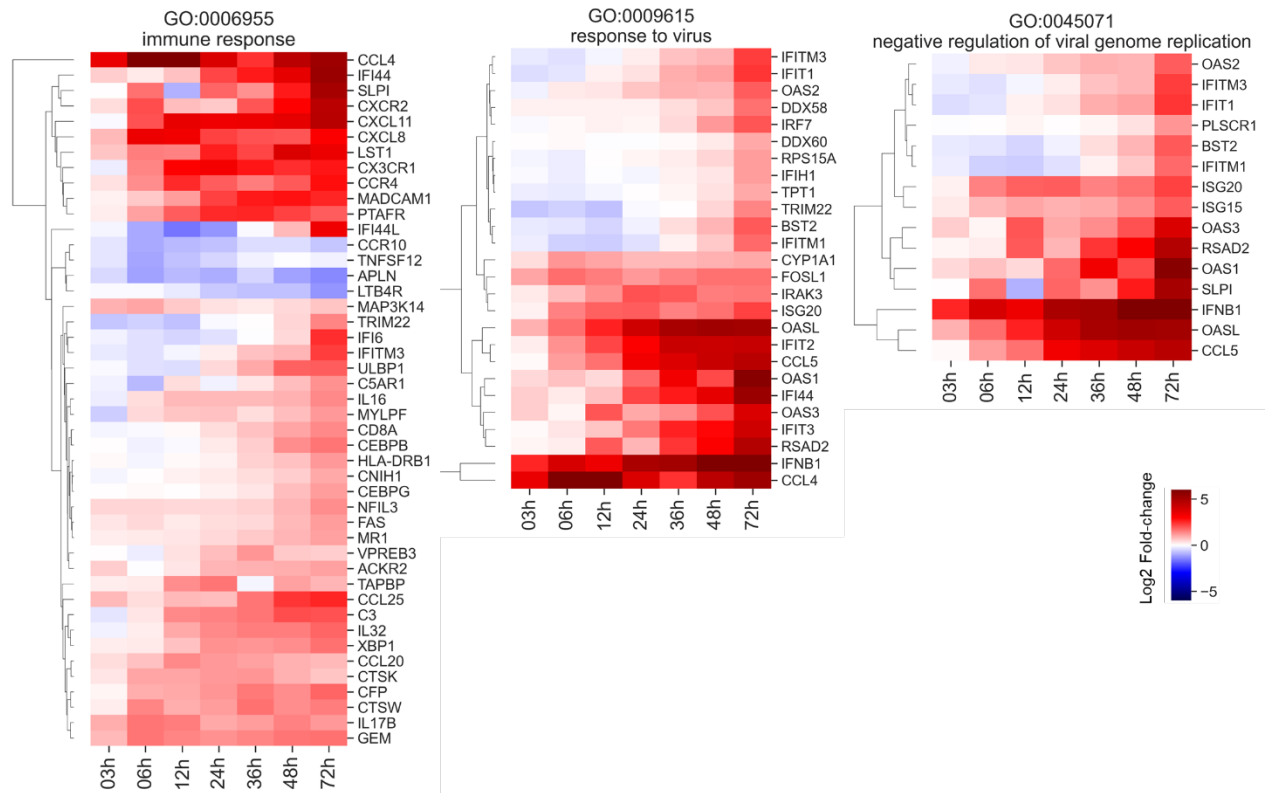


Figure S4. Hierarchical clustering of genes involved in response to virus and 6 related biological processes significantly enriched in GO enrichment analysis. Heatmap was ordered by time points. All listed genes were DEGs that passed the selection criteria described in Figure S1B. Expression fold-changes were calculated compared with pre-transfection time point 0h. The fold changes were indicated in log2 scale by a hue of red/white/blue color pad that represents up-regulated/basal/down-regulated expression level.

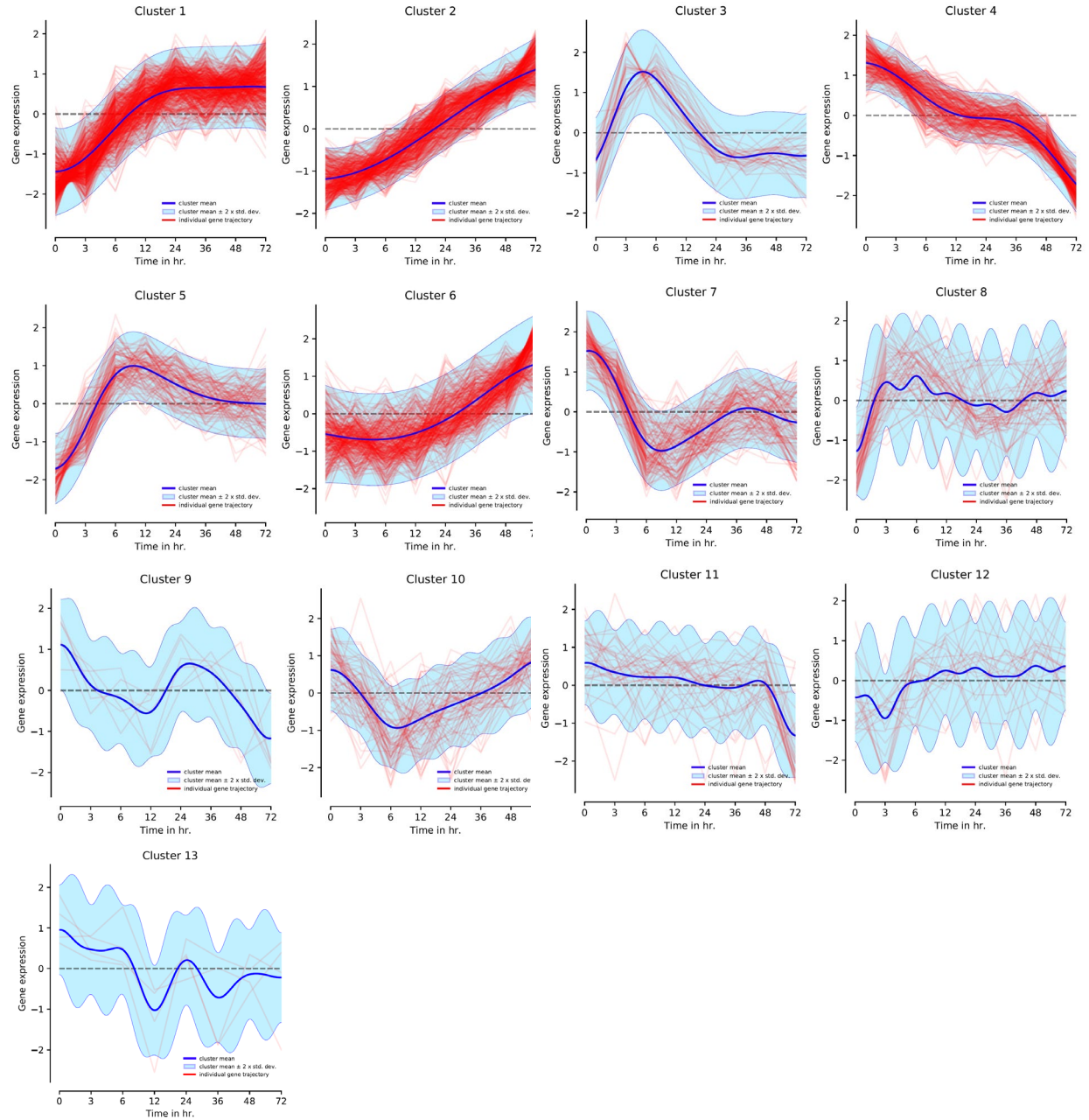


Figure S5. Time series analysis using DPGP mixture model classified 1,850 DEGs into 13 distinct clusters. The x-axis represents sampling time points. The y-axis represents standardized log2 fold change standardized to zero mean and unit variance across 8 time points and triplicates. Red lines indicate standardized expression level for individual genes. The dark blue line represents the mean of clustered genes while the blue shade area is the range of  $\pm 2$  standard deviation.



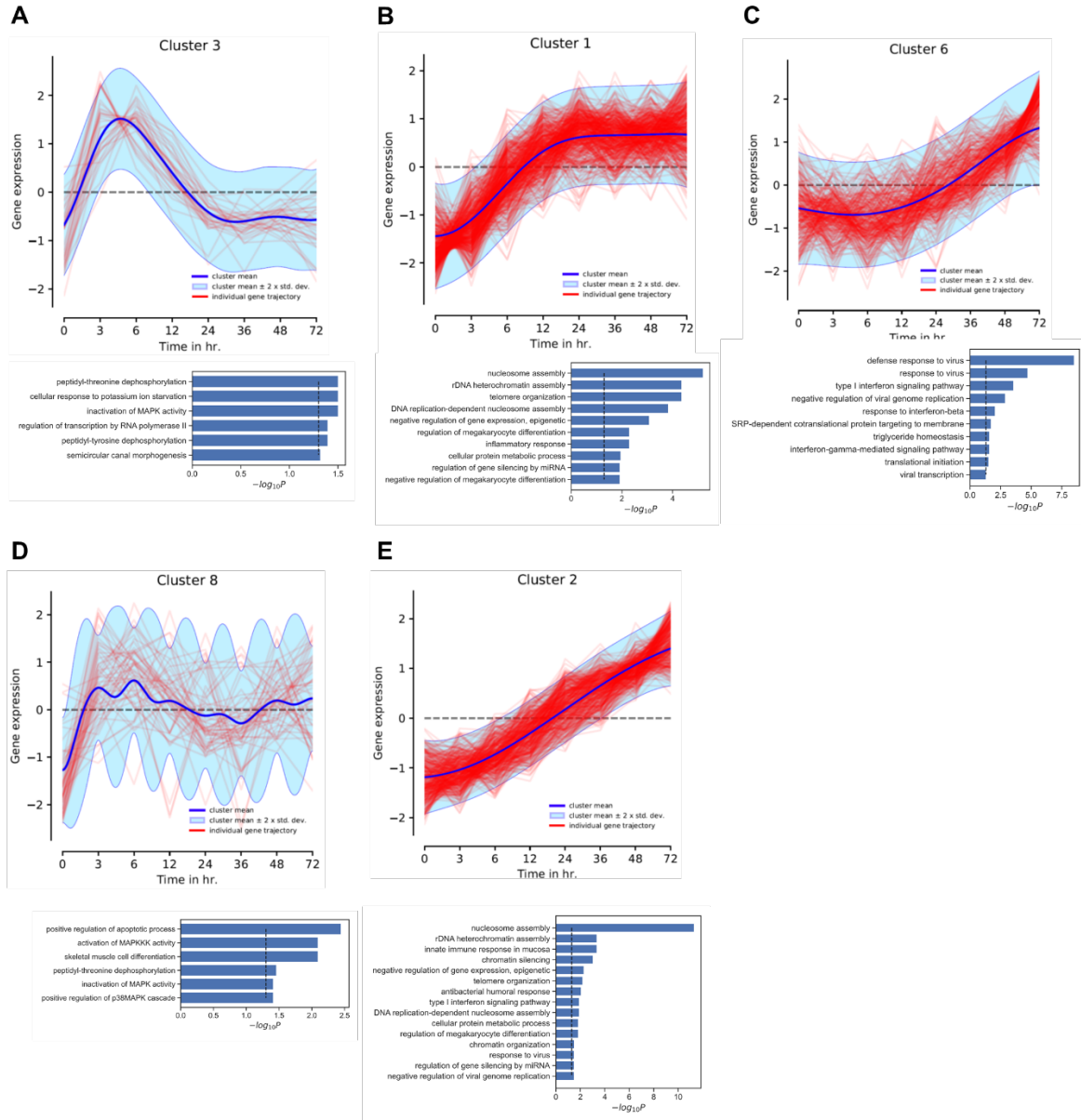


Figure S6. DPGP identified clusters and the significantly enriched gene ontologies identified from clustered DEGs. The x-axis shows the adjusted p-value (FDR) after multiple testing correction using Benjamini-Hochberg procedure. GO terms were significant when FDR < 0.05, black dashed line.

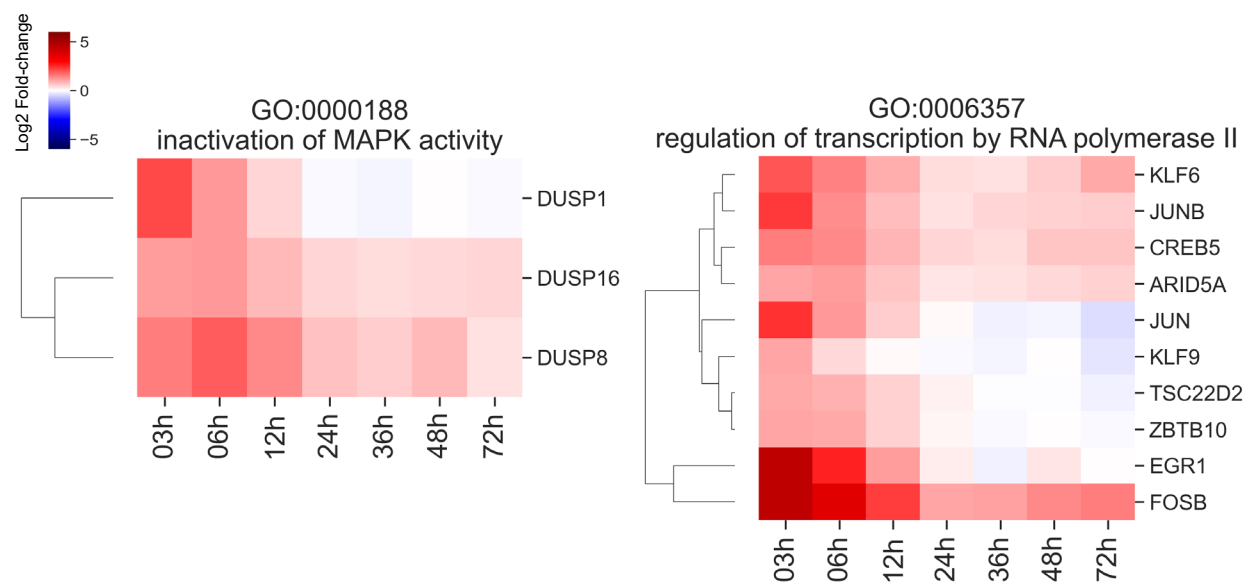


Figure S7. Hierarchical clustering of genes involved in the Early responding cluster and 2 related biological processes significantly enriched in GO enrichment analysis. Heatmap was ordered by time points. All listed genes were DEGs that passed the selection criteria described in Figure S1B. Expression fold-changes were calculated compared with pre-transfection time point 0h. The fold changes were indicated in log2 scale by a hue of red/white/blue color pad that represents up-regulated/basal/down-regulated expression level.

## References:

1. McDowell, I.C., Manandhar, D., Vockley, C.M., Schmid, A.K., Reddy, T.E., and Engelhardt, B.E. (2018). Clustering gene expression time series data using an infinite Gaussian process mixture model. *PLoS Comput Biol* 14, e1005896. [10.1371/journal.pcbi.1005896](https://doi.org/10.1371/journal.pcbi.1005896).
2. Klopfenstein, D.V., Zhang, L., Pedersen, B.S., Ramirez, F., Warwick Vesztrocy, A., Naldi, A., Mungall, C.J., Yunes, J.M., Botvinnik, O., Weigel, M., Dampier, W., et al. (2018). GOATOOLS: A Python library for Gene Ontology analyses. *Sci Rep* 8, 10872. [10.1038/s41598-018-28948-z](https://doi.org/10.1038/s41598-018-28948-z).
3. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 34, 525-527. [10.1038/nbt.3519](https://doi.org/10.1038/nbt.3519).