Molecular Therapy
Methods & Clinical Development
Original Article

# E2A, VA RNA I, and L4-22k adenoviral helper genes are sufficient for AAV production in HEK293 cells

Jiten Doshi,[1,2,3] Emma Couto,[1] Jillian Staiti,[1] Luk H. Vandenberghe,[1,2,3,4] and Nerea Zabaleta[1,2,3,4]

[1]Schepens Eye Research Institute, Mass Eye and Ear Infirmary, Boston, MA, USA; [2]Harvard Medical School, Boston, MA, USA; [3]Broad Institute of MIT and Harvard, Boston, MA, USA

**The replication-defective adeno-associated virus (AAV) is extensively utilized as a research tool or vector for gene therapy. The production process of AAV remains intricate, expensive, and mechanistically underexplored. With the aim of enhancing AAV manufacturing efficiencies in mammalian cells, we revisited the questions and optimization surrounding the requirement of the various adenoviral helper genes in enabling AAV production. First, we refined the minimal set of adenoviral genes in HEK293 AAV production to *E2A*, *L4-22K/33K*, and *VA RNA I*. These findings challenge the previously accepted necessity of adenoviral *E4orf6* in AAV production. In addition, we identified *L4-22K* genes as crucial helpers for AAV production. Next, a revised minimal adenoviral helper plasmid comprising *E2A*, *L4-22K*, and *VA RNA I* genes was designed and demonstrated to yield high titer and potent AAV preps in HEK293 transient transfection. Lastly, stable packaging cells harboring inducible *E2A* and *L4-22K* genes were shown to maintain AAV production yields comparable to transient transfection over a culture period of ~10 weeks. Combined, these findings further our understanding of adenoviral helper function in mammalian AAV production and provide novel plasmid and cell-line reagents with an improved safety profile for potential broad applicability in the research and gene therapy community.**

## INTRODUCTION

Adeno-associated viruses (AAVs), which belong to the genus *Dependoparvovirus* within the *Parvoviridae* family, are non-enveloped single-stranded DNA (ssDNA) viruses of 4.7-kb genome consisting of packaging *rep* and functional *cap* genes flanked by inverted terminal repeat (ITR) sequences. Wild-type AAV infections are not considered disease causing. AAV-based vectors have proved favorable for gene therapy/gene delivery applications due to their advantageous features, such as eliciting mild immune responses, efficient targeting of specific tissues, and facilitating long-term gene expression. Wild-type AAVs require co-infection with a helper virus—either adenovirus or herpes simplex virus—to complete its viral replication cycle, providing an additional safety feature to the AAV viral vectors. An increasing number of AAV-based therapies has received regulatory approval based on an active slate of programs at the various phases of clinical testing.[1,2] The demand for clinical-grade AAV vectors is therefore expected to continue to increase to support several products.[3,4]

One of the predominant methodologies to generate AAV is through transient transfection of mammalian human embryonic kidney (HEK) 293 cells with multiple plasmids. Briefly, (1) packaging *rep* and structural *cap* genes, (2) recombinant AAV genome flanked by ITRs, and (3) adenovirus helper genes are encoded on multiple plasmids to minimize the chance for a replication-competent viral sequence to be reconstituted. Next, these plasmids are co-transfected in HEK293 cells to then be harvested at some point post transfection, generally after 72 h. HEK293 cells were originally derived in 1973 from the kidney of an aborted human embryo, following transformation with a 4-kb adenoviral DNA fragment containing the E1A/B and protein IX genes, which integrated into chromosome 19.[5] These cells are considered pseudo-triploid, with a higher copy number (5–6) for the region of chromosome 19 that harbors the adenoviral DNA.[6] Although initially thought to originate from kidney tissue, transcriptomic profiling suggests that HEK293 cells may actually be derived from adrenal or neuronal tissues.[6] Their rapid growth rate, ease of culture, high transfection efficiency, and integrated adenoviral sequences make them particularly well suited for efficient AAV production. This method is relatively easy to establish and versatile for research phases where various components of the product are being varied and optimized. However, high transfection costs, requirement of costly Good Manufacturing Practices (GMP)-grade plasmid DNA, limited scalability, and challenging reproducibility of this complex process at scale limit the ability to meet the rising demand of AAV vectors.[7,8] This is further complicated by challenges related to the generation of large proportions of empty or only partially filled capsid

Received 29 July 2024; accepted 8 November 2024;
https://doi.org/10.1016/j.omtm.2024.101376.

[4]These authors contributed equally

**Correspondence:** Luk H. Vandenberghe, Schepens Eye Research Institute, Mass Eye and Ear Infirmary, Boston, MA, USA.
**E-mail:** luk_vandenberghe@meei.harvard.edu
**Correspondence:** Nerea Zabaleta, Schepens Eye Research Institute, Mass Eye and Ear Infirmary, Boston, MA, USA.
**E-mail:** nerea_zabaletalasarte@meei.harvard.edu

**Table 1. Role of adenoviral genes in AAV production**

| Gene | Role(s) in AAV production |
|---|---|
| *E1A* | master regulator of gene expression of both AAV and adenovirus genes[26–28] |
| *E1B-E4orf6* | proteasomal degradation of p53 and MRN complex[31–36]<br>export of viral mRNA and inhibition of export of cellular mRNA from the nucleus to cytoplasm[15,16,37,38] |
| *E2A* | ssDBP that is involved in the AAV genome replication[29,30] |
| *VA RNA I* | inhibition of protein kinase R, thereby improving the translation of viral particles[19–21]<br>impacts miRNA biogenesis pathway<br>substrate for RNA interference[41–44]<br>translational control[19] |
| *L4-22K/33K* | amplification of Rep-Cap sequences in stable packaging cells[45]<br>mRNA splicing and export to cytoplasm[49–51]<br>genome packaging? |
| *E4orf6* | enhances AAV transduction through promotion of second strand DNA synthesis during the replication process[15,17,39]<br>degradation of AAV5 Capsid and Rep52 proteins[40] |

(virus particles containing no or partial ssDNA genome) and other impurities within a preparation.[9] Reduction in manufacturing related costs, e.g., by eliminating the requirement of GMP-grade plasmid DNA and using stable cell lines, will have a direct impact on reducing the overall therapy cost.

In the antibody manufacturing space, high yields are made possible at large scales by using stable cell lines.[10] Prior to the generation of a stable cell line, we sought to revisit the question as to what the minimal adenoviral helper genes are to support the production of AAV vectors. This effort was inspired by the hypothesis that the current synthetic biology armamentarium may overcome some methodological limitations in the studies from over 20 years ago that defined *E1A/B*, *E2A*, *E4orf6*, and *VA RNA I* genes to be sufficient for AAV production.[11,12] More recently, *L4-22K/33K* adenovirus genes were (referred to from hereon as *22K/33K*) identified to be indispensable for AAV production.[13,14] Plus, the currently used adenovirus helper plasmids (e.g., pDeltaF6 or pALD-X80) consist of many more genes and non-coding sequences (including regulatory elements) than the previously identified minimal gene set and it is unknown if they play an important role in AAV production. Defining the minimal set of adenovirus helper genes required to support HEK293-based AAV production enables the development of optimized plasmid and cell-line reagents, aiming to reduce the cost of AAV manufacturing. Furthermore, they may enhance the safety profile of AAV products by eliminating unnecessary components from the manufacturing process and possible residual contamination in the final product.

In a mechanistic sense, AAV production consists of three main steps: AAV genome replication, capsid biosynthesis, and genome packaging. Processes such as viral mRNA synthesis and export to cytoplasm,[15–17] viral protein synthesis,[18,19] suppression of cellular antiviral responses,[20–23] secretion of assembled viral particles[24,25] also indirectly contribute to production. AAV Rep and Cap proteins are central to the production steps. Rep proteins (Rep78, Rep68, Rep52, Rep40) play a key role in AAV genome replication and packaging. Cap proteins, VP1/2/3, make up the capsid shell and interact with Rep proteins during genome packaging. Apart from AAV genes, several cellular and adenoviral genes (Table 1) play important roles in one or more of these steps. In the case of adenoviral genes, *E1A* is a master regulator of gene expression of both AAV and adenovirus genes.[26–28] E2A, being a ssDNA-binding protein (ssDBP), is involved in the AAV genome replication.[29,30] E1B associates with E4orf6 and other cellular proteins to form a ubiquitin protein ligase complex that directs proteasomal degradation of p53 and MRN complex.[31–36] The complex is demonstrated to facilitate export of viral mRNA and inhibit export of cellular mRNA from the nucleus to cytoplasm.[15,16,37,38] In addition, it is shown to enhance AAV transduction through promotion of second-strand DNA synthesis during the replication process.[15,17,39] E4orf6 is also shown to be involved in direct degradation of AAV5 capsid and Rep52 proteins.[40] VA RNA I is mainly involved in inhibition of Protein Krepinase R (recognizes dsRNA and exerts antiviral effects by inhibiting translation and production of new viral particles), thereby improving the translation of viral particles.[19–21] VA RNA I is also found to impact the miRNA biogenesis pathway and itself serves as a substrate for RNA interference.[41–44] Mechanisms through which contributive nature of recently known adenoviral *22K* and *33K* genes in AAV production is not clear.[13,14] A recent study has highlighted the importance of 22K and 33K proteins in the amplification of *rep/cap* sequences in stable packaging cells.[45] Both 22K and 33K in the context of adenovirus life cycle have multiple functions. 33K is an alternative splicing factor, while 22K gene is involved in the regulation of expression of early and late adenoviral genes.[46,47] Both 22K and 33K proteins are implicated in adenovirus genome packaging and exhibit interactions with various adenoviral proteins.[46–48]

Stable cell lines for AAV production requires integration of multiple genes. Previous attempts to integrate many of the helper genes have resulted in significant yield reduction, as overexpression of these genes can be toxic to the cells and their integration into the genome causes cellular instability.[52–54] Following the identification of the minimal gene set, tight and inducible control of helper gene expression could overcome these limitations.[53]

Here, we identify the previously unknown 22K/33K adenovirus genes as indispensable for AAV production, confirming findings by Adsero et al.. and Johari et al.[13,14] Additionally, we show *E4orf6* can be omitted in HEK293 AAV production while maintaining high yields of potent vector. We propose, based on our findings, a refined minimal adenoviral helper gene set for HEK293-based AAV vector production: *E1A/B*, *E2A*, *VA RNA I*, and *22K/33K*. A newly designed helper plasmid, JD171, which encodes this minimal set, produces AAV vectors at comparable levels to pDeltaF6 and pALD-X80 in yield, packaging, and transduction efficiency. Lastly, we show that the stable cells harboring inducible *E2A* and *22K/33K* genes successfully produce AAV vectors at yields comparable to that achieved by

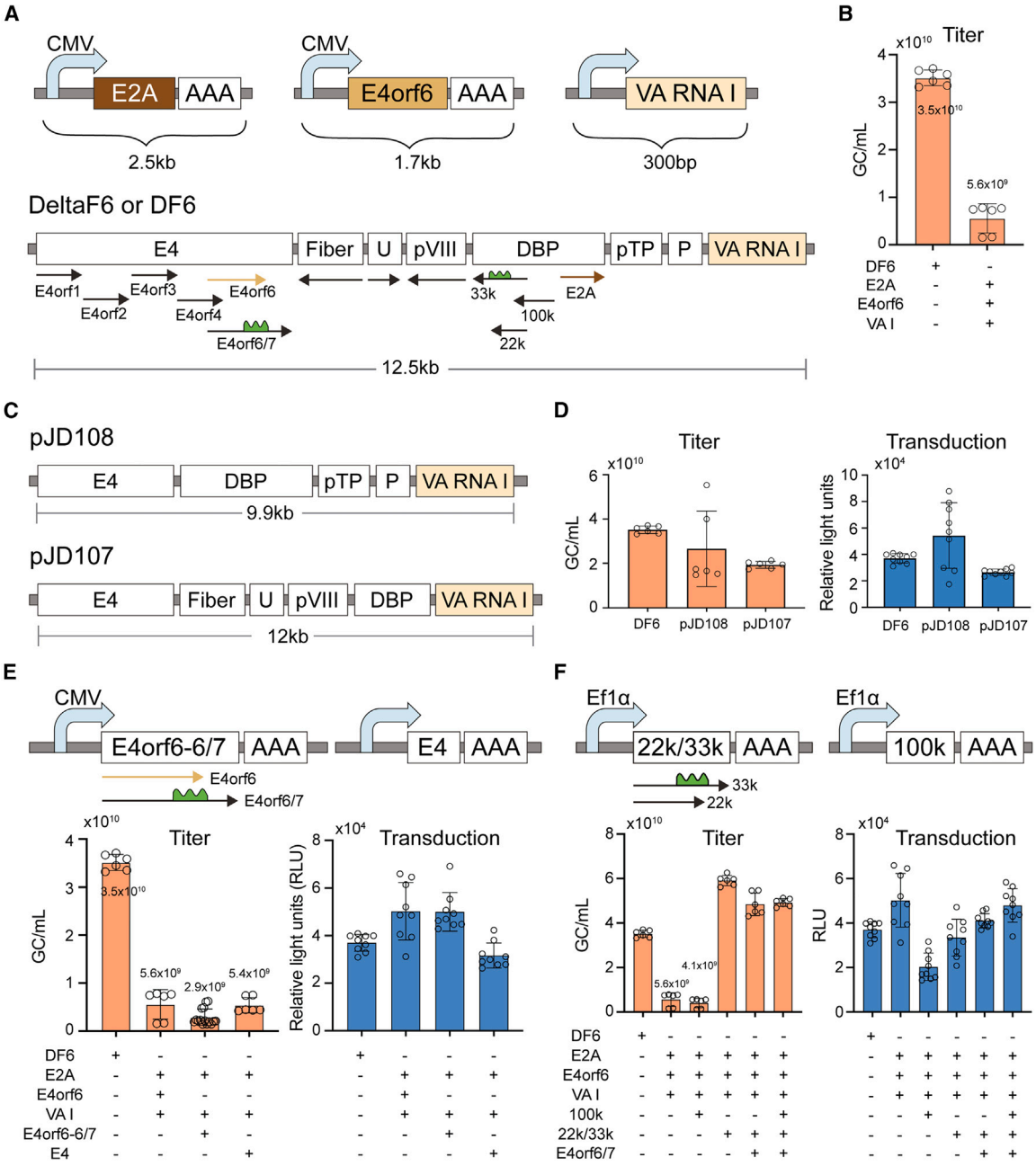



**Figure 1. Assessing the impact of adenovirus helper genes on AAV production resulted in the identification of essentiality of *22K/33K* genes**

(A) Schematic depicting the essential *E2A*, *E4orf6*, and *VA RNA I* adenoviral genes with their sizes (top) and commonly used helper plasmid DF6 (bottom) with its open reading frames (ORFs). Black arrows indicate the direction and length of protein coding sequence in the plasmid. Sliding waves in green indicate intronic sequences. Colored arrows and rectangle are used to denote ORFs of the known essential genes *E2A*, *E4orf6*, and *VA RNA I*. U, U exon; DBP, DNA-binding protein; pTP, precursor terminal protein; P, protease. (B) The bar chart shows AAV9 titer in GC/mL measured via qPCR when using DF6 vs. a combination of *E2A*, *E4orf6*, and *VA RNA I* helper genes. (C) Schematic of helper plasmids pJD108 (top) and pJD107 (bottom), both derived from DF6. (D) Bar charts show AAV9 titer in GC/mL (left) and transduction efficiency at a multiplicity of infection (MOI) of $10^4$ GC/cell in relative light units (RLU) (right) for helper plasmids DF6, pJD108, and pJD107. (E) Top, schematic of *E4orf6-6/7* and *E4* genes. E4orf6/7 (black) and E4orf6 (colored) ORFs are indicated as arrows. Sliding waves in green indicate intronic sequences. Bottom, bar charts showing AAV9 titer in GC/mL (left) and transduction efficiency at a MOI of $10^4$ GC/cell in RLU (right) for a combination of helper genes/plasmids, presence indicated by + sign and absence indicated by − sign. (F) Top, schematic of *22K/33K* and *100K* genes. 22K and 33K ORFs are indicated as arrows. Sliding waves in green indicate intronic sequences. Bottom, bar charts showing AAV9 titer in GC/mL (left) and transduction efficiency in RLU at a MOI of $10^4$ GC/cell for a combination of helper genes/plasmids, presence indicated by + sign and absence indicated by − sign. Each bar in (B), (D), (E), and (F) represents mean ± SD of technical triplicates wherein readouts were obtained from $n = 3$ biological replicates. Some bars

*(legend continued on next page)*

transient transfection. Overall, by deconstructing adenovirus helper requirements, identifying the minimal gene set, and providing a proof of concept for stably integrating helpers for AAV production, our work takes a step toward reducing the cost and complexity of AAV manufacturing, while furthering our understanding of the biological mechanism of helper function in the HEK293 system.

## RESULTS

### *E2A*, *E4orf6*, and *VA RNA I* helper genes are not sufficient for AAV production

Prior studies determined the essential adenoviral components for AAV vector production in HEK293 to comprise of *E2A*, *E4orf6*, and *VA RNA I* genes (in addition to the integrated, constitutively expressed *E1A/B*). These genes, including constitutive regulatory elements, make up a total of approximately 4.5 kb in size (Figure 1A). However, commonly utilized helpers like pDeltaF6 (DF6) and pALD-X80 (ALD-X80) harbor adenoviral sequences in addition to these genes, spanning ∼12.5 and ∼15 kb respectively (Figures 1A and S1A). The surplus content in DF6 includes complete coding sequences of 11 additional adenoviral genes and partial sequences of five others. These supplementary regulatory and coding adenoviral sequences confer no known benefit toward AAV production. Moreover, the presence of unwanted viral sequence could pose a safety risk and increases transfection costs.

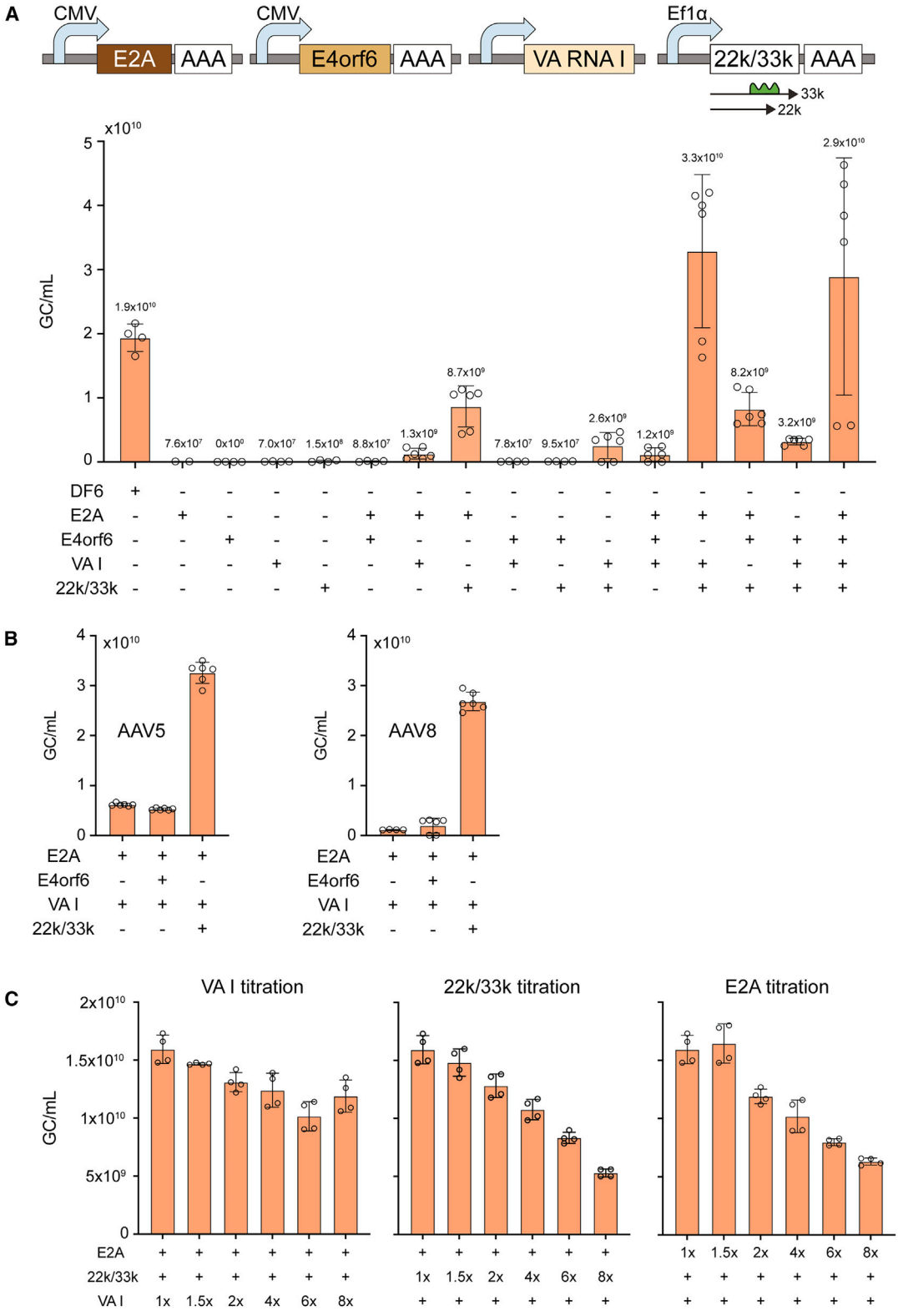
To understand their significance in the context of AAV production, we cloned *E2A*, *E4orf6*, and *VA RNA I* genes in individual plasmids. The expression of *E2A* and *E4orf6* was driven by the constitutive cytomegalovirus (CMV) promoter while *VA RNA I* expression was driven by its wild-type RNA polymerase III promoter (Figure 1A). We then compared titers of AAV preparations (24-well plate; production scale 500 μL) via qPCR; transfections had either equimolar amounts of the three individual genes or the commonly used helper plasmid DF6. Alongside the helper plasmid(s), a recombinant AAV genome containing the *luciferase* gene (pJD36: ITR-CMV-Luciferase-ITR) and a plasmid encoding for AAV Rep2 and AAV capsid9 serotype (pJD64) were transfected. qPCR assay using a luciferase probe (Table S1) showed that the AAV9 vector titer represented as genome copies (GC) per mL (GC/mL) was ∼7-fold lower when using *E2A*, *E4orf6*, and *VA RNA I* genes than when DF6 was transfected (Figure 1B). Luciferase intensity, represented in relative light units (RLU), measured 48 h post transduction of freshly seeded HEK293 cells at multiplicity of infection (MOI) of $10^4$ GC/cell with the two AAV9 preps, did not reveal any significant difference (Figure S1B). The expression of the helper genes *E2A*, *E4orf6*, and *VA RNA I* was detectable at the transcript level from the individual plasmid at a similar, but not identical level, compared to the DF6 helper (Figure S1C). This study suggested that E2A, E4orf6, and VA RNA I enable production of AAV9 vectors but with low efficiency and/or require a very specific stoichiometry for optimal AAV production.

### 22K/33K gene is indispensable and *E4orf6* is dispensable for AAV production

In the light of previous results, we hypothesized that the additional sequence present in the helper plasmid could further potentiate AAV production. Analysis of all open reading frames within the helper plasmids DF6 and ALD-X80 (Figures 1A and S1A) was performed and two additional helper plasmids, pJD107 and pJD108, were generated from DF6; pJD108 was devoid of *fiber*, *U exon*, and *pVIII* genes and pJD107 was devoid of partial sequences of protease and percursor terminal protein (pTP) (Figure 1C). Interestingly, either one of the novel helper plasmids rescued titers and retained transduction efficiency of vector preparations relative to the DF6 benchmark (Figure 1D). We concluded from this that *fiber*, *U exon*, *pVIII* genes, or partial sequences of protease and pTP genes did not have any impact on production, and, therefore, they were considered non-essential.

Then, we explored the relevance of the E4 adenoviral region that contains the essential *E4orf6* gene and five other open reading frames (ORFs). We supplied plasmids encoding either just E4orf6, the combination of E4orf6 + E4orf6/7, or the complete E4 region alongside the other two base helper genes (*E2A*, *VA RNA I*). Presence of these supplementary E4 elements did not elevate AAV9 vector titers to levels achieved with DF6 helper (Figure 1E), despite maintaining transduction capability (Figure 1E). Next, the critical *E2A* gene operates within the adenovirus DNA-binding protein (DBP) coding region, also housing 100K, 22K, and 33K genes (Figure 1F). The 22K and 33K proteins exhibit an overlapping N terminus but divergent C termini, leading us to evaluate them in combination. While 100K was unable to rescue production titer supporting previous observations,[13,45,55] 22K/33K proteins recovered titers to levels observed with DF6 without impacting transduction efficiency (Figure 1F). When the *22K* and *33K* genes were cloned into separate plasmids and evaluated individually, both were found to contribute to enhancing AAV9 production titers, with the *22K* gene exhibiting a more substantial influence compared to the *33K* gene (Figure S1D). The addition of E4orf6/7 and/or 100K over 22K/33K failed to further improve titers (Figure 1F). In this setting, 23K and VA RNA II were also shown not to further enhance AAV production (Figures S1D and S1E). Negligible yields in control samples lacking helpers or AAV components confirmed assay specificity (Figure S1F). Collectively, these data underscore the importance of the 22K/33K adenoviral gene products in AAV production. In contrast, the adenoviral *fiber*, *U exon*, *pVIII*, *pTP*, *E4orfs* (1, 2, 3, 4, 6/7), *100K*, *23K* (protease), and *VA RNA II* genes confer no production benefit.

are shown with respective plotted values. Titer and transduction value when using DF6 (or E2A, E4orf6, VA RNA I) is identical in the bar charts because all the conditions in this figure were tested in a single experiment. All the samples had pJD64 (Rep2-Cap9) and pJD36 (AAV genome: CMV-Luciferase) alongside helper genes/plasmids during transfection. CMV, human cytomegalovirus promoter; Ef1 $\alpha$, human elongation factor 1 alpha promoter; AAA, polyadenylation signal. Titers in GC/mL were measured via qPCR. See "materials and methods" section for details on qPCR titer assay and luciferase transduction efficiency assay.




*(legend on next page)*

This newly established role for 22K/33K was recently also described by Adsero et al. and Johari et al.[13,14] and we further sought to reassess whether the previously defined minimal set of adenoviral helper genes for AAV production was held up in the 22K/33K context. AAV9 preparations were made with all possible genetic combinations drawn from *E2A*, *E4orf6*, *VA RNA I*, and the newly implicated *22K/33K* genes. As expected, single genes result in negligible production titer (Figure 2A). Some of the two gene subsets resulted in meaningful titers, with the combined *E2A* and *22K/33K* titers reaching almost to 50% of titers achieved with DF6 (Figure 2A). However, other combinations yielded much less, indicating the relative importance of these two helper genes (Figures 2A and S2A). One combination of three genes, namely *E2A*, *VA RNA I*, and *22K/33K*, resulted in a meaningful titer that in fact was comparable to DF6 and the combination of all four genes. The rest of the three gene combinations resulted in titers 2- to 10-fold lower than DF6 (Figure S2A). As expected, the entire four-gene set resulted in titers as observed previously. Overall, the data showed that E4orf6 is dispensable for production in a context where E2A, VA RNA I, and 22K/33K are present. Additionally, we observed that a DF6 plasmid variant devoid of any E4 sequence produced a comparable titer to DF6 (Figure S2B), further supporting the limited role of E4orf6 in this context. Absence of either E2A or 22K/33K resulted in a greater than 10-fold reduction of titer, indicative of an essential role for these co-factors. In comparison, the absence of VA RNA I resulted in a 2- to 3-fold drop in titers (Figure S2A). Additionally, our data do not illustrate any impact on the quality of the preparation as assessed by transduction efficiency *in vitro* (Figure S2C). Lastly, we also examined the minimal adenoviral helper genes set for AAV serotypes 5 and 8. Titers of both serotypes were highest when *E2A*, *VA RNA I*, and *22K/33K* genes were supplied (Figure 2B). Similar to AAV9, presence of E4orf6 had no impact on titers of AAV5 and AAV8 (compare E2A + VA RNA I + E4orf6 vs. E2A + VA RNA I). Notably, for AAV5, titers with *E2A* and *VA RNA I* genes were ∼6-fold lower than when using *E2A*, *VA RNA I*, and *22K/33K* genes. For AAV9, this difference in titer was ∼17-fold (Figure 2A). Our results thus indicate that E2A, VA RNA I, and 22K/33K are required and sufficient to provide adenoviral helper functions during transfection-based AAV production.

We next performed titrations of *E2A*, *22K/33K*, and *VA RNA I* quantities to determine optimal ratios relative to the Rep-Cap plasmid for maximizing yields. Our titration data showed that increasing molar amounts of each of the helper genes did not result in any increase in titers (Figure 2C). By titration of <1× molar values, we observed that 0.5× molar amounts each of *E2A*, *22K/33K*, and *VA RNA I* helper genes conferred peak yields (Figures 2C and S3A). Based on these data, we concluded that 0.5× each of *E2A* and *22K/33K* resulted in a meaningful benefit in titers. However, upon combining the optimal values, AAV9 vector titers remained consistently similar to when 1× molar equivalence of each of *E2A*, *22K/33K*, and *VA RNA I* genes was used, indicating a possible stochiometric relationship of these helper genes in AAV production (Figure S3B). The titration was not pursued further as we explored placing these helper genes together in a single plasmid.

### Helper plasmids optimized for safety, size, and yield show scalable production efficiency

Applying the insights from earlier studies, we sought to optimize helper plasmids that incorporated the minimal adenoviral genes set our data supports (pJD171 or JD171: *E2A*, *VA RNA I*, and *22K/33K*) or the four adenoviral genes set (pJD150 or JD150: *E2A*, *VA RNA I*, *E4orf6*, and *22K/33K*). Plasmid designs were considered with either a constitutive CMV or an EF1$\alpha$ promoter that drives *E4orf6*, *E2A*, and *22K/33K* expression (Figure 3A). The bovine growth hormone (bGH) or SV40 polyadenylation signal was included for transcription termination of *E2A* and *22K/33K*, respectively, while the native signal was retained for *E4orf6* gene. *VA RNA I* gene utilized its native wild-type signals. The new helper plasmids JD150 (four genes) and JD171 (three genes) are respectively 5.2 and 7.2 kb shorter in adenoviral sequence than DF6 (Figure 3A). Production titers (Figure 3B) and transduction efficiency, tested in HEK293 (Figure S3C) and HeLa (Figure S3D) cells, of the new helpers JD150 and JD171 remained comparable to helpers DF6 and ALD-X80 in a 24-well production format. Next, we up-scaled the production to 15-cm dish (three independent runs, production scale: 18 mL) and the vectors were purified using a static binding assay with AAV9 POROS resin resulting in the equal capture of full and empty capsids. Titers and transduction capability still remained equivalent for the preps produced using the three helpers (DF6, JD150, JD171) after up-scaling (Figure 3C; Tables 2, S3E, and S3F). Capsids per mL determined by ELISA displayed a similar trend (Figure 3D). Transmission electron microscopy images of the purified AAV9 preps displayed highly similar morphology between the different conditions (Figure 3E). Overall, the new helper plasmids that were optimized for size and

**Figure 2. Identification of minimal adenoviral helper gene set and their optimal amounts for production of AAV vectors**

(A) Top, schematic depicting *E2A*, *E4orf6*, *VA RNA I*, and *22K/33K* adenoviral genes with constitutive regulatory sequences. Bottom, bar chart displaying titer in GC/mL for the combinations of *E2A*, *E4orf6*, *VA RNA I*, and *22K/33K* helper genes used in transfection for AAV9 production. Presence of a gene is indicated by + sign and absence is indicated by − sign. Titer in GC/mL with DeltaF6 (DF6) plasmid used as a standard value. Bars are shown with respective plotted values. (B) Bar chart displaying titer in GC/mL for the combinations of *E2A*, *E4orf6*, *VA RNA I*, and *22K/33K* helper genes used in transfection for AAV5 (left) and AAV8 (right) production. Presence of a gene is indicated by + sign and absence is indicated by − sign. Rep2-Cap5 and Rep2-Cap8 plasmids were respectively used for production of AAV5 and AAV8 serotypes. AAV genome plasmid (pJD36 CMV-Luciferase) was used for production of both the serotypes. (C) Titration of *VA RNA I* (left), *22K/33K* (middle), and *E2A* (right) helper genes. Bar charts represent titer in GC/mL measured via qPCR. The molar equivalent to pJD64 (Rep2-Cap9) plasmid is shown in the bottom of the charts for the respective helper genes. + sign indicates presence and is equivalent to 1× molar equivalence to pJD64 (Rep2-Cap9) plasmid. The bar corresponding to 1× molar equivalence of *E2A*, *22K/33K*, and *VA RNA I* is the same in all the charts because the titrations were performed in a single experiment. Each bar in (A)–(C) represents mean ± SD of technical duplicates wherein readouts were obtained from biological replicates ($n$ = 2 or 3). All samples in (A) and (C) had pJD64 (Rep2-Cap9) and pJD36 (AAV genome: CMV-Luciferase) alongside helper genes/plasmids during transfection. Titers in GC/mL were measured via qPCR. See "materials and methods" section for details on qPCR titer assay.

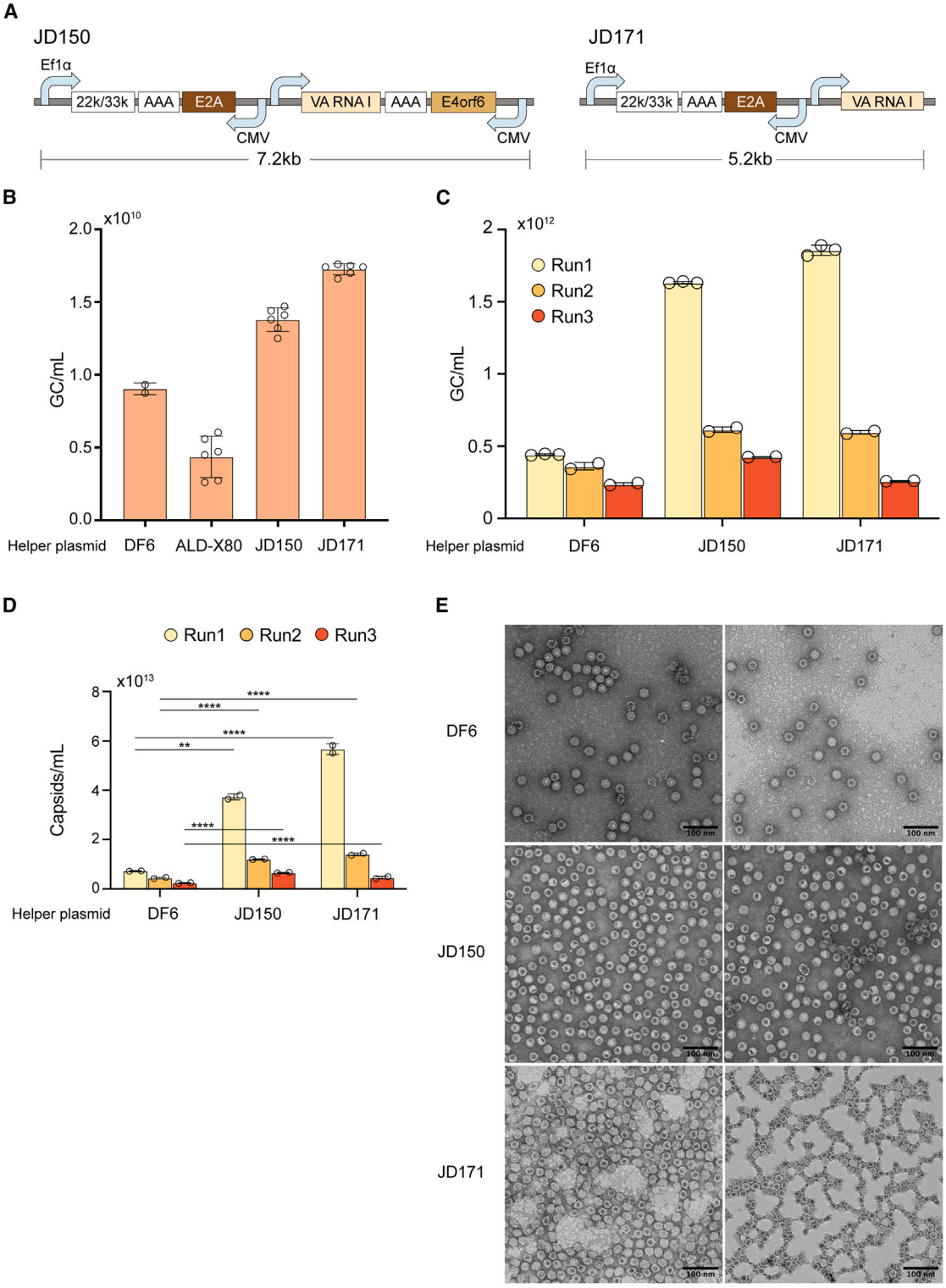



**Figure 3. Optimized helper plasmids for AAV production**

(A) Schematic depicting two helper plasmids JD150 and JD171 with its genes and regulatory sequences. CMV, human cytomegalovirus promoter; Ef1 $\alpha$, human elongation factor 1 alpha promoter; AAA, polyadenylation signal. (B) Bar chart displaying titer in GC/mL measured via qPCR for AAV9 vector produced using DF6, ALD-X80, JD150, and

*(legend continued on next page)*

the number of helper genes demonstrated equivalence to DF6 or ALD-X80 regarding key AAV production attributes including titer, transduction efficiency, and total particles. In summary, these newer helper systems provide reduced size, eliminating non-essential adenoviral sequences and lower transfection requirements, conferring potential opportunities for cost reduction in manufacturing.

### Production efficiency with the new helpers is independent of serotype or vector genome

The versatility of the optimized helper plasmids was evaluated through vector production of multiple serotypes (AAV5, 6, 8, and 11) and recombinant genomes alongside comparative benchmarks. AAV11 vector production was done with a CMV-SARS-CoV-2 delta spike[56] recombinant genome, while CMV-Luciferase (pJD36) was used for AAV5, 6, and 8. Comparable vector titers were achieved across serotypes when employing either of the novel helper constructs JD150 or JD171, or the conventionally utilized DF6 helper (Figures 4A and S4A). Transduction efficiency in HeLa cells remained equivalent for AAV6 vector preps made using DF6 and JD171 helpers (Figure S4B). Given the equivalent performance, further experiments were done with the JD171 helper since it contains only the minimal set of helper genes. Increase in production scale (500 mL) and substitution with an alternate AAV genome (a reporter cassette comprising ITR-flanked CMV-mScarlet or eGFP) resulted in AAV9 vectors titer to be 1.5- to 2.2-fold lower when using JD171 helper than conventional helper ALD-X80 (Figures 4B and 4C). Improved fold change in titers of preps made with JD171 and ALD-X80 helpers between the two runs can be attributed to the increase made in the copy number ratio of JD171 helper to 1 from 0.8 relative to Rep2-Cap9 plasmid (Figures 4B and 4C). Quantitative analysis of transduction efficiency by flow cytometry of AAV9 preps containing mScarlet transgene further exhibited less than 2-fold variances across the panel (Figure S4C). The data demonstrated the capacity to support vector production when using these novel helper constructs across AAV serotypes and distinct recombinant genome configurations.

### AAV production using stable cell lines containing inducible *E2A*, *22K/33K* helper genes

Next, we investigated whether the need for transfection of the helper genes could be overcome entirely through the generation of stable cell lines. Since constitutive expression of some helper genes may impart toxicity and thus impact AAV production and cell viability,[52–54] we configured *E2A* and *22K/33K* under a bi-directional tet-responsive promoter to achieve inducible control of their expression when supplemented with tetracycline transactivator (tTA) protein. Since *VA RNA I* is transcribed by RNA polymerase III promoter, generating an inducible expression system poses challenges. Our previous results have indicated that the presence of *VA RNA I* leads to 2- to 3-fold increase in production titers, and its absence consequently does not completely abolish production. Thus, our helper construct maintained *VA RNA I* under constitutive expression while placing the other adenoviral genes under inducible control (Figure 5A). Prior to genomic integration, we confirmed that the novel helper construct stringently regulates AAV yields using transient transfection; in the absence of tTA induction, negligible titers were observed, while addition increased titers 40- to 55-fold, to levels comparable to DF6 (Figure 5A).

The above-mentioned helper genes design was transferred to a piggybac transfer vector and integrated into HEK293 cells using the piggybac transposon (Figure S5A). Alongside the helper genes, the transfer vector also contained the mCitrine fluorescent gene cassette under the control of inducible erythromycin repressible promoter (ETR).[57] Following integration, positive integrants in polyclonal and monoclonal form were sorted based on mCitrine fluorescence, activated by transfection of the erythromycin transactivator (ET) (Figures S5B and S5C). AAV production using the stable cells required transfection of tTA inducer, Rep2-Cap9, and AAV ITR-flanked genome plasmids. Out of the polyclonal and monoclonal lines tested at 3–4 weeks post sorting, highest titers were observed with 3M4 and 3M8 monoclonal variants (Figure S5D). Overall, titers from monoclonal lines were 6- to 20-fold lower when compared to titer obtained from transient transfection of HEK293 cells. Re-testing the monoclonal variants 3M4 and 3M8 in 3–4 weeks from the initial production run revealed 3- to 10-fold decline in titers (Figure S5E). We hypothesized that the decline in titers could be attributed to cellular instability induced by the constitutive nature of *VA RNA I* expression. Furthermore, titers were elevated in the absence of tTA inducer, hinting at leakiness in expression of *E2A*, *22K/33K* genes.

Given the potential toxicity of the constitutive expression of *VA RNA I*, we decided to only integrate the *E2A* and *22K/33K* genes. Integration and subsequent sorting of monoclonal and polyclonal populations were undertaken as described previously (Figures 5B

JD171 helpers in 24-well plate (500 μL production scale). Each bar represents mean ± SD of technical triplicates wherein readouts were obtained from biological replicates ($n$ = 3). Exception being DF6 value, which is represented by mean ± SD of technical duplicate wherein readout was obtained from a single biological replicate. Bar chart displaying titer in GC/mL (C) and capsids/mL measured by ELISA (D) for AAV9 vector produced using DF6, JD150, and JD171 helpers in three independent runs (runs 1, 2, and 3) performed in 15-cm dish. The reduced yields in runs 2 and 3 across the three helpers is attributed to the change in the number of cells seeded and the amount of AAV9 POROS resin utilized during downstream processing. Refer to "materials and methods" section for details. Each bar in (C) and (D) represents mean ± SD of technical replicates ($n$ = 2 or 3) wherein readouts were obtained from a single biological replicate. Statistical analysis in (D) was performed using one-way ANOVA with Tukey's *post hoc* test (adjusted p-value <0.0001 or 0.003). (E) Transmission electron microscopy images of AAV9 vectors produced using DF6, JD150, and JD171 helpers. Full particles appear as dark, solid circles or spheres with a distinct boundary while empty particles appear as light or less intense circles or spheres. Scale bar, 100 nm as indicated. Magnification 98,000×. All samples in (B) and (C) had pJD64 (Rep2-Cap9) and pJD36 (AAV genome: CMV-Luciferase) alongside helper plasmid in triple transfection. Helper plasmid used for production is denoted below the bars. Titers in GC/mL were measured via qPCR. See "materials and methods" section for details on qPCR titer assay and capsid ELISA.

**Table 2. Genome titer and capsids/mL data for AAVs made using helpers DF6, JD150, and JD171**

| | Run1 | | | Run2 | | | Run3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | DF6 | JD150 | JD171 | DF6 | JD150 | JD171 | DF6 | JD150 | JD171 |
| Titer (GC/mL) | 4.43E+11 | 1.63E+12 | 1.85E+12 | 3.62E+11 | 6.15E+11 | 5.97E+11 | 2.38E+11 | 4.25E+11 | 2.59E+11 |
| Capsids/mL | 7.25E+12 | 3.73E+13 | 5.67E+13 | 4.43E+12 | 1.20E+13 | 1.40E+13 | 2.42E+12 | 6.57E+12 | 4.57E+12 |

and S6A). The viability of the polyclonal cells remained in the range of 77%–98% over 2 months of culturing (Figure S6B). In these stable cells, AAV production required transfection of tTA inducer, VA RNA I, Rep2-Cap9, and AAV ITR-flanked genome plasmids. Among the eight monoclonal lines tested at 3–4 weeks post sorting, the 2M4 and 2M8 lines yielded the highest vector titers (Figure 5C). Subsequent re-testing of the monoclonal lines 2M4 and 2M8 at 3–4 weeks after the initial production run revealed that titers in the absence of tTA inducer (VA RNA I + Rep2-Cap9) were 10- and 2-fold lower, respectively, than in the presence of tTA (tTA + VA RNA I + Rep2-Cap9) (Figure 5D). As expected, the production of AAV9 vectors was reduced, by 2- to 3-fold, when *VA RNA I* was absent compared to when it was present. Since 2M4 had negligible titers in the absence of tTA, it was further re-tested in the third production round at ∼3 weeks after the previous run. High titer values were observed in the presence of tTA and 100-fold lower titer in its absence (Figure 5E). Importantly, the titer in the presence of tTA (tTA + VA RNA I + Rep2-Cap9 condition) was comparable to titer value obtained when the same elements, in the form of JD171 helper, were transiently transfected in HEK293 cells. Copy number analysis using ddPCR on 2M4 cells revealed approximately 20 copies integrated into the genome (Table 3). Overall, these data point to the applicability of integrating helper genes, specifically *E2A* and *22K/33K*, in the cells to produce AAV vectors. Genomic context dependency could be the reason that resulted in observation of meaningful AAV titers in the absence of tTA in certain monoclonal lines. To conclude, we demonstrated successful integration of inducible forms of *22K/33K* and *E2A* genes and its precise control of expression consequently controlled AAV production. The monoclonal line, 2M4, showed robust and stable production titer, comparable to transient transfection over a reasonable culture duration of ∼10 weeks. Thus, our work serves as a steppingstone for developing a truly stable producer cell line for AAV production.

## DISCUSSION

The production of viral vectors can be visualized as an isolated synthetic system wherein the cellular factory provides the basic resources. Development of an isolated system requires replacement of native and complex regulation elements with synthetic, controllable elements. Our work takes a step in this direction, uncoupling parts of the complex regulation pattern of adenovirus, AAV, and cellular genes governing AAV production.

In this study, firstly, we systematically delineated the minimal set of adenoviral helper genes necessary for production of AAV vectors within the HEK293 mammalian cell production system. Our data indicated that the adenoviral *fiber*, *U exon*, *pVIII*, *pTP*, *100K*, *23K* (protease), *VA RNA II*, *E4orf1-4*, and *E4orf6/7* genes do not play a role in AAV production while confirming the essential role of *22K/33K* genes. Of the two, *22K* had substantial impact on AAV titers. Combinatorial testing of the final four helper genes (*E2A*, *22K/33K*, *E4orf6*, and *VA RNA I*) indicated (1) *E2A* and *22K/33K* synergistically enhance titers making both genes "critical" for production; (2) *VA RNA I* acts as a "booster," increasing titers by 2- to 3-fold; and (3) absence of *E4orf6* had no impact on AAV preps, contrary to the previous belief of its essentiality. With this knowledge, we generated new adenoviral helper plasmids, harboring only the three helper genes, leading to a plasmid size benefit and eliminating unnecessary adenoviral components of possible safety concern. The novel helper plasmids are comparable in quantitative and qualitative metrics, scalable as well as equivalent at producing different AAV serotypes carrying diverse recombinant AAV genomes. Additionally, we have generated a helper cell line that stably expresses two essential adenovirus helper genes in an inducible manner. This cell line shows stability across passages, minimal leakiness in the absence of inducer, and production efficiency on par with triple transfection. These improvements may therefore reduce manufacturing costs associated with gene-therapy treatments as well as increasing safety standards.

Analysis of AAV vector preps through next-generation sequencing has been shown to possess around 0.5%–2% of contaminant adenoviral helper sequences.[58–60] In addition, structural adenoviral proteins such as fiber protein can elicit immune reactions and its absence in AAV preps is vital.[11] Our work here shows that the transfection of only three helper genes (*E2A*, *VA RNA I*, and *22K/33K*) mimics the production efficacy of larger adenoviral helper plasmids. The new helper devoid of E4 region and other unwanted adenoviral genes (*fiber*, *U exon*, *pVIII*, *pTP*, *100K*, *23K* [protease], *VA RNA II*) improves safety by reducing the presence of the contaminant adenoviral DNA sequences and proteins otherwise found in the vector preparations.[13,59]

It is still unclear what the exact role of 22K/33K is in AAV production. Previous observations from studies about 22K include (1) 22K mutant adenoviruses result in empty capsids suggesting its role in genome replication/packaging,[46] (2) 22K has been implicated in regulation of expression of late genes in adenovirus,[47] and (3) 22K has been shown to be involved in the splicing of adenoviral genes and *rep*.[45,49,50] Recently, 22K was shown to associate with human coilin protein in Cajal bodies and is involved in mRNA processing and export to cytoplasm.[51] The latter function is analogous to the role of E1B-E4orf6 complex. Therefore, the mechanism of action of 22K

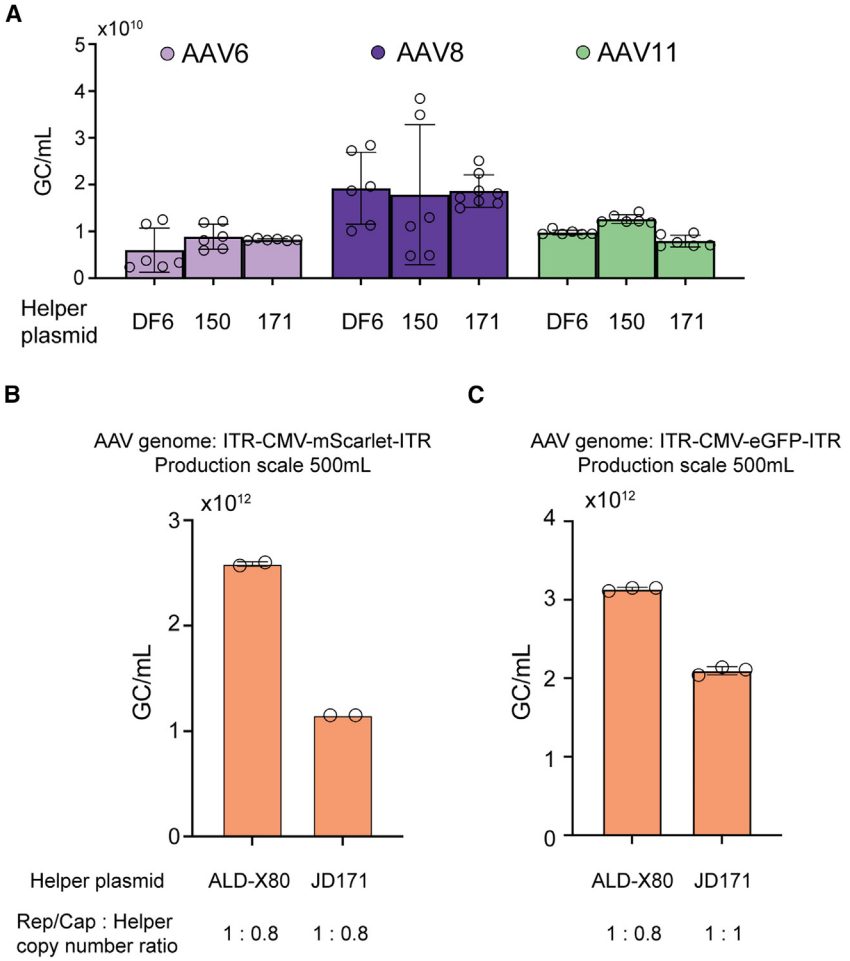



**Figure 4. AAV production using next-generation helper plasmids is independent of serotype, genome, and scale**

(A) Bar chart displaying titer in GC/mL for AAV6, AAV8, and AAV11 vectors produced using DF6, JD150, and JD171 helpers in a 24-well plate. AAV genome containing CMV-SARS-CoV-2 delta spike was used in case of AAV11 and AAV genome containing CMV-luciferase (pJD36) was used in case of AAV6 and AAV8 production. Rep2-Cap6, Rep2-Cap8, and Rep2-Cap11 were utilized, respectively, for the production of AAV6, AAV8, and AAV11 vectors. Each bar represents mean ± SD of technical duplicates wherein readouts were obtained from three biological replicates. (B and C) Bar chart displaying titer in GC/mL for AAV9 vector produced using ALD-X80 and JD171 helpers in HYPERflask at a scale of 500 mL. Helper plasmid used for production and copy number ratio of Rep-Cap to helper plasmid is displayed below the bars. AAV genome containing CMV-mScarlet (pJD34) was used for AAV9 vector production in (B). AAV genome containing CMV-eGFP was used for AAV9 vector production in (C). Each bar in (B) and (C) represents mean ± SD of technical replicates ($n$ = 2 or 3) wherein readouts were obtained from a single biological replicate. Samples in (B) and (C) used pJD64 (Rep2-Cap9) plasmid in triple transfection. Helper plasmid used for production is denoted below the bars. Titers in GC/mL were measured via qPCR. See "materials and methods" section for details on qPCR titer assay.

is likely related to certain redundant functions like mRNA processing and transport while also having some essential role in genome replication/packaging. E4orf6-E1B complex is shown to be crucial in increasing AAV second-strand synthesis and thus improving AAV transduction.[15,39,61,62] However, as per our knowledge, the same enhancing effect of E4orf6 on AAV replication has not been studied in the context of AAV production. Despite its role in transduction, it could be that the function of E4orf6 in the context of AAV production is either redundant or remains absent altogether.

In the development of gene therapy products, one of the key challenges is to reduce the manufacturing costs associated with the production process. Producing plasmid DNA that meets GMP standards and the use of transfection reagent contribute significantly to these expenses.[8] Our newly designed helper, JD171, approximately 10 kb smaller than ALD-X80, demonstrates potential benefits for large-scale manufacturing processes by requiring lower quantities to achieve equivalent production yields. Our data show that a 500-mL production run (HYPERflask, 1,720-cm$^2$ surface area) utilized 300 μg of JD171 helper compared to 520 μg of ALD-X80. Considering the GMP-grade plasmid cost of $100,000/μg and extrapolating to a 200-L production run (iCELLis, 133m$^2$), the use of JD171 helper would result in estimated savings of ∼$17,000.[8] Moreover, the reduced requirement of transfection reagent (FectoVIR) with JD171 helper contributes to additional savings of ∼$22,690. Collectively, these factors lead to potential savings of ∼$39,700 per production run. Furthermore, the new helper plasmid can seamlessly replace the current helpers in the existing transient transfection-based production systems utilizing HEK293 cells, facilitating an efficient transition. This improvement in helper design not only offers significant cost savings but also maintains compatibility with established production protocols.

Our work also serves as a proof of concept for using stable helper cell lines for AAV production. We demonstrated that the monoclonal cell line 2M4 containing integrated *E2A* and *22K/33K* genes had AAV9 titers comparable to what is observed with transient transfection method in HEK293 cells. Stable integration of helper sequences in DF6 or ALD-X80 was not performed for comparison because previous studies have already outlined the cellular toxicity and instability caused by constitutive expression of the helper genes.[52–54] In some monoclonal lines, we observed meaningful AAV9 titers in the absence of tTA inducer, likely attributed to the leakiness of expression of *E2A* and *22K/33K* genes. Targeted integration at safe-harbor locations in the genome could possibly remove the observed leakiness by mitigating positional effects and enhance tight control of gene expression,

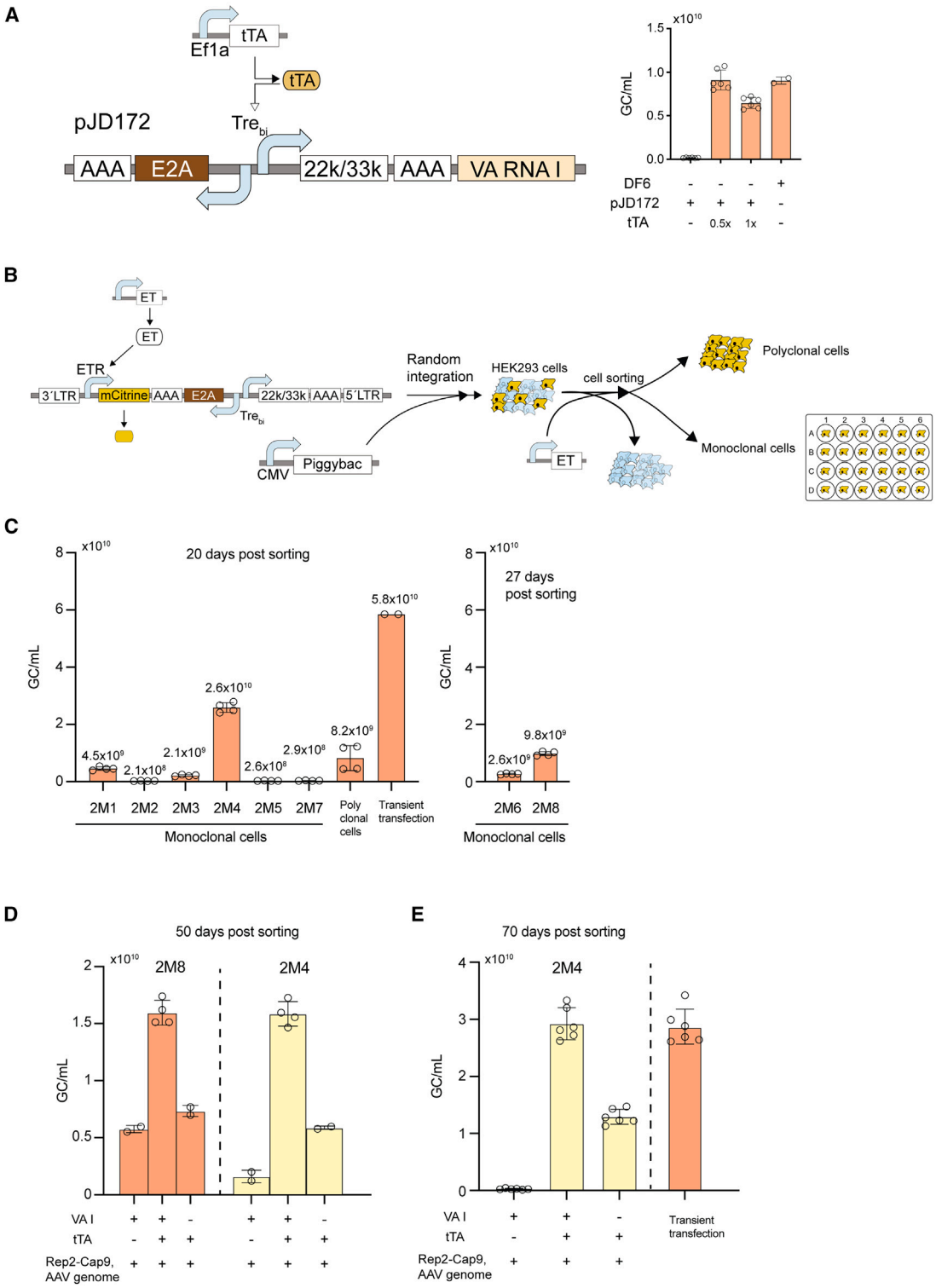



**Figure 5. Stable cell lines with inducible *E2A* and *22K/33K* helper genes for AAV production**

(A) Schematic on the left displaying bi-directional tet inducible promoter ($Tre_{bi}$) driving expression of *E2A* and *22K/33K* genes. The inducible promoter is activated by tetracycline transactivator (tTA) provided as a separate plasmid. Bar chart (right) displaying titer in GC/mL for AAV9 produced in 24-well plate using either DF6 or inducible pJD172 helper plasmid. Presence of a gene/helper plasmid is indicated by + sign and absence is indicated by − sign. For tTA, 0.5× or 1× molar equivalent to pJD64

*(legend continued on next page)*

which could potentially aid in further improving vector production. The *E2A* gene is known to be well tolerated in host cells,[63] which likely contributed to the observed stability of the viral vector titers in the stable cell lines; however, for *22K/33K* genes, it has not yet been established to what extent it is tolerated in a stable expression context in HEK293. *VA RNA I* helper gene remains crucial as it provides 2- to 3-fold improvement in vector titers. The constitutive expression of *VA RNA I* could be a potential source contributing to the diminishing viral vector yields observed when all three helper genes (*E2A*, *22K/33K*, *VA RNA I*) were integrated into the stable cell lines, and improvements in the design strategy may help overcome this issue. Overall, inducible AAV production using stable helper cell lines with optimal copy number of helper genes could help eliminate the need for transfection of at least one plasmid.

To conclude, we have developed a new helper plasmid and a truly stable cell line harboring two of the three helper genes. Both these outcomes serve as a crucial starting point for streamlining AAV vector production and reducing manufacturing costs, thereby make gene therapies more accessible and affordable.

## MATERIALS AND METHODS

### Cloning

For the various kits used, manufacturer's instructions were followed unless indicated otherwise. Plasmids and its cloning procedures are listed in Table S1. DNA amplification was performed using Phusion High Fidelity DNA Polymerase (NEB, catalog no. [Cat#]M0530). De-salted primers/oligonucleotides (Table S1) were ordered from Azenta or IDT. De-salted gene fragments/synthetic DNA sequences/gBlocks (Table S3) were ordered from IDT. Digestion fragments were purified using MinElute PCR purification kit (Qiagen, Cat#28004). Gel extraction and purification was performed using Zymo Gel DNA recovery kit (Zymo, Cat#D4001). Restriction digestion was performed at temperatures recommended by the manufacturer. Ligation reaction was performed at room temperature (RT) using T4 DNA ligase (NEB, Cat#M0202). Mix and Go DH5alpha *E. coli* transformation kit (Zymo, Cat#T3001) or NEB Stable cells (NEB, Cat#C3040H) or NEB DH5alpha cells (NEB, Cat#C2987H) were used for transformation. Screening of positive clones was performed using colony PCR with Quick-Load Taq 2× Master Mix (NEB, Cat#M0271L). Plasmid isolation from positive clones was performed using Zymo PURE Plasmid mini-prep kit (Zymo, Cat#D4210). All the plasmids were verified by either using Sanger sequencing service provided by Azenta or whole-plasmid sequencing service provided by MGH CCIB DNA Core. Transformed bacteria were cultured in LB broth (RPI, Cat#L24040-2000) supplemented with appropriate antibiotics (carbenicillin disodium salt 100 mg/mL [Boston bioproduct, Cat#P760] or kanamycin 50 mg/mL [Boston Bioproducts, Cat#P-810] or chloramphenicol [RPI, Cat#C61000-25]). 1× NEB HiFi DNA assembly mix (NEB, Cat#E2621L) was used to perform Gibson assembly at 50°C for 45 min by mixing vector (50 ng) and insert(s) (3–5 M equivalent) in a final volume of 10 μL. Oligo cloning required phosphorylation and annealing of oligonucleotides prior to ligation with the backbone fragment. Phosphorylation of oligonucleotides was performed by adding together 1.5 μL of oligonucleotide (100 μM), 2.5 μL of 10× polynucleotide kinase (PNK) buffer, 2.5 μL of ATP (10 mM), 0.5μL of T4 PNK enzyme (10 U/mL) (NEB, Cat#M0201S), and 18 μL of $ddH_2O$ followed by incubation at 37°C for 30 min. Annealing was performed by mixing 25 μL each of the phosphorylated oligonucleotide and then incubating in a thermocycler at 95°C for 2 min followed by a decrease of 0.5°C every minute for the next 44 min. 1 μL of 1:10 or 1:20 diluted (with $ddH_2O$) annealed oligonucleotides was used for ligation reaction. Golden gate assembly was carried out using BsaI or BbsI restriction enzymes following publicly available protocols.

### Cell culture

All the experiments were performed in HEK293 cells (Lonza), cultured at 37°C, 5% $CO_2$, in 0.2 μm (Corning, Cat#431097) of filtered DMEM (Corning, Cat#10-013-CV) supplemented with 10% FBS (R&D Systems, Cat#S12450) and 1% penicillin/streptomycin solution (Corning, Cat#30-002CI). Splitting was performed every 3–4 days using 0.25% trypsin-EDTA (Corning, Cat#25-053-CI). Cell cultures were propagated for at most 2 months before being replaced by a fresh cell stock.

### AAV production in 24-well plate

24-well plates (Corning, Cat#3524) were pre-coated with poly-L-lysine (Sigma, Cat#P4707) following manufacturer's instructions.

(Rep2-Cap9) plasmid was used for activation. Each bar represents mean ± SD of technical duplicates wherein readouts were obtained from biological replicates ($n = 3$). (B) Schematic showing the process of random integration of transfer vector in the HEK293 genome using the piggybac transposon system followed by fluorescence-based sorting. Piggybac transfer vector containing inducible *E2A* and *22K/33K* helper genes and inducible mCitrine fluorescent protein shown on left. Random integration in HEK293 cells by transfection of transfer vector and piggybac transposon. mCitrine expression is under the control of erythromycin repressible promoter (ETR) activated by erythromycin-dependent transactivator (ET). Cells were transfected with ET plasmid to activate mCitrine expression 48 h before sorting. Monoclonal and polyclonal cells were isolated through sorting. (C) Bar charts show titer in GC/mL for AAV9 produced in 24-well plate scale using monoclonal (2M1-2M8), polyclonal stable helper cells as well as with HEK293 transient transfection. VA RNA I, Rep2-Cap9, AAV genome, and tTA inducer plasmids were transfected in monoclonal and polyclonal cells for production. Analyzing inducibility and stability of vector production in monoclonal 2M4 and 2M8 cells after 50 days post sorting in (D) and in monoclonal 2M4 after 70 days post sorting in (E). Bar charts in (D) and (E) show titer in GC/mL for AAV9 vectors produced in 24-well plate. In (D) and (E), presence of VA RNA I, Rep2-Cap9, AAV genome, and tTA inducer plasmids during transfection is indicated by + sign and absence is indicated by − sign. JD171 helper, Rep2-Cap9, and AAV genome were transfected in case of production in transient transfection condition in (C) and (E). AAV genome composed of CMV-Luciferase (pJD36). Bars in (C) and (D) represent mean ± SD of technical duplicates wherein readouts were obtained from either biological duplicates or a single replicate. Bars in (E) represent mean ± SD of technical duplicates wherein readouts were obtained from $n = 3$ biological replicates. AAA, polyadenylation signal; $Tre_{bi}$, bi-directional tetracycline responsive promoter. Titers in GC/mL were measured via qPCR. See "materials and methods" section for details on stable integration, sorting, and qPCR titer assay.

**Table 3. Copy number analysis of 2M4 monoclonal cells using ddPCR**

| | Copies per diploid genome | | |
|---|---|---|---|
| Sample | Replicate 1 | Replicate 2 | Replicate 3 |
| 2M4 monoclonal cells | 20.33 | 19.28 | 19.91 |
| HEK293 cells | 0.06 | 0.04 | 0.03 |
| No template | 0.01 | 0.00 | 0.01 |

HEK293 cells were counted using TC20 automated cell counter (Bio-Rad, Cat#1450102) and seeded on pre-coated plates at a density of $1.65 \times 10^5$ cells per well in 500 µL of complete DMEM (supplemented with 10% FBS and 1% penicillin/streptomycin) for achieving a confluency of 85%–95% at the time of transfection. The seeded plates were incubated at 37°C, 5% $CO_2$ for ∼24 h. A ratio of 1:1.375 was used of DNA (µg) to PEImax (µg) (Polysciences, Cat#24765-1) for transfection. Appropriate volume of PEImax was taken and serum-free DMEM was used to make final volume of 50 µL (PEImax-DMEM mix). Appropriate plasmid DNA amounts were taken and mixed with serum-free DMEM to make final volume of 50 µL (DMEM-DNA mix). The DMEM-DNA mix was incubated for 5 min before mixing with appropriate amount of PEImax-DMEM mix. The resulting DNA-PEImax mix was then incubated at RT for 10–20 min. Medium was gently removed from the wells of 24-well plate, which was supplied with 400 µL of serum-free DMEM (supplemented with 1% penicillin/streptomycin). 100 µL of DNA-PEImax mix was added dropwise to the respective wells. Volumetric errors were accounted by preparing 1.5× of PEImax-DMEM mix and 1.25× of DMEM-DNA mix.

The Rep-Cap and helper plasmids used to generate the data represented in Figures 1, 2, S1, and S2 were combined in equal molar ratios prior to being transfected. Ratio for Rep-Cap:AAV genome plasmid was 1:1.25. For reference, 258 ng each of Rep2-Cap9 (pJD64) plasmid and 520 ng of DF6 helper plasmid (Addgene plasmid #112867, gift from James Wilson University of Pennsylvania) pertaining to $\sim 3.3 \times 10^{10}$ copies and 258 ng of AAV genome plasmid (pJD36, CMV-luciferase) pertaining to $\sim 4.2 \times 10^{10}$ copies were used for transfection of a well in 24-well plate. Refer to Table S1 for plasmid sizes and for plasmids utilized in each sample of each figure. New helpers (JD150 and JD171) were utilized in molar ratio of 1:0.8:1.25 corresponding to Rep2-CapX (6/8/9/5/11):helper:AAV genome for data in Figures 3B and 4A. pALD-X80 plasmid (Aldevron, Cat#5017-10) was utilized as helper at equimolar ratio to Rep-Cap plasmid for data shown in Figure 3B. No-helper, no-transfection, and no-cap-gene control conditions were tested in a couple of experiments (Figure S1F).

The plasmids pJD64 (Rep2-Cap9), pJD36 (AAV genome), pJD33 (Ef1α-tTA), and pJD80 (VA RNA I) were utilized in transfections as required for generating the data represented in Figures 5A, 5C–5E, S5D, and S5E. The inducer plasmid pJD33 was used at 0.5× and 1× molar equivalence to Rep2-Cap9 plasmid in Figure 5A. The inducer plasmid pJD33 was used at 0.75× molar equivalence to Rep2-Cap9 plasmid for data in Figures 5C–5E, S5D, and S5E.

For crude AAV harvest/purification, three cycles of freezing (−80°C for 30 min) and thawing (37°C for 30 min) of plate/samples were performed 72 h post transfection. Then, the plates/samples were centrifuged at >3,000 × *g* for 20 min and the supernatant was used for downstream analysis (qPCR, capsid ELISA, and luciferase transduction assay).

### AAV production in 15-cm dish

AAVs were produced in a 15-cm dish for obtaining data corresponding to Figure 3C. 15-cm dishes (Corning, Cat#430599) were pre-coated with poly-L-lysine (Sigma, Cat#P4707) by following manufacturer's instructions. HEK293 cells were seeded on coated 15-cm dishes at a density of $1.7 \times 10^7$ cells (Figure 3C run 2 and run 3) or $1.9 \times 10^7$ cells (Figure 3C run 1) per dish in 16–18 mL of DMEM (supplemented with 10% FBS and 1% penicillin/streptomycin) for achieving a confluency of 85%–95% at the time of transfection. The seeded dishes were incubated at 37°C, 5% $CO_2$ for ∼24 h. A ratio of 1:1.375 was used of DNA (µg) to PEImax (µg). The plasmid quantities utilized in 24-well-plate transfections were increased by a proportional scaling factor of 80 (equivalent increase in surface area) to ascertain the amounts required for 15-cm-dish transfections. New helpers (JD150 and JD171) were utilized in molar ratio of 1:0.8:1.25 corresponding to Rep2-Cap9:helper:AAV genome. DF6 helper plasmid was mixed in equimolar fashion with Rep2-Cap9 and AAV genome plasmids prior to transfection. Appropriate amounts of plasmids used for each transfection were mixed together and serum-free DMEM was used to make final volume of 1 mL (DNA-DMEM mix). Plasmid amounts used were as follows: 20.64 µg each of pJD64 (Rep2-Cap9) and pJD36 (AAV genome: CMV-luciferase) and either of the three helper plasmids (DF6 41.6 µg or pJD150 23.04 µg, pJD171 18.8 µg). Appropriate volume of PEImax was taken and serum-free DMEM was used to make final volume of 1 mL (PEImax-DMEM mix). PEImax-DMEM mix was incubated at RT for 5 min. Following incubation, the mix was added to DNA-DMEM mix to make DNA-PEImax mix and incubated for 15–20 min. During the incubation period, medium was gently removed from the 15-cm dishes and supplied with fresh 16–18 mL of serum-free DMEM (supplemented with 1% penicillin/streptomycin). Following incubation, DNA-PEImax mix was added dropwise to the cells. The dishes were processed for AAV purification via static resin binding/capture 72 h post transfection as described below.

### AAV purification via static resin binding/capture

AAVs were isolated using POROS AAV9 resin (Thermo, Cat#A27354) through a static binding process, as previously described.[64] Briefly, 72 h post transfection, cells were lysed and treated with nucleases. The clarified lysate was then incubated with AAV9 POROS resin to capture AAV particles, which were then eluted from the resin and concentrated using an Amicon cassette. The detailed protocol is as follows: 72 h post transfection, the supernatant medium was decanted in a 15-mL falcon tube and mixed with Triton X-100 (5 µL per 1 mL of lysate) (Alfa Aesar, Cat#A16046), RNase A (4.2 µg per 1mL of lysate) (Millipore, Cat#70856), Turbonuclease (25 U per 1 mL of lysate) (Accelagen, 250U/µL Cat#N0301M), and 1% final concentration of Pluronic F-68 (Sigma; Cat#P556). The contents were then poured back into the

production dish and incubated 37°C for 1 h with occasional mixing. The supernatant containing the lysate was then transferred from the production dish to a 50-mL falcon tube. The production dish was further washed with PBS (230 μL per 1 mL of lysate) (Corning, Cat#21-031-CM) and added to the rest of the lysate. The lysate was centrifuged at 4,000 × *g* at 4°C for 30 min followed by filtration of supernatant with 0.2-μm filter (Corning, Cat#431229).

50 μL (Figure 3C Run1) or 15 μL (Figure 3C Run2 and Run3) of POROS AAV9 resin (Thermo, Cat#A27354) was conditioned by performing three washes with 4 mL of 0.1 M NaCl (Fisher Scientific Cat#BP358-1). The resin was pelleted after each wash and supernatant discarded by pulse centrifugation (2,500 × *g* for 2 min 30 s). The resin was then re-suspended in 1 mL of PBS and 0.001% Pluronic F68 for equilibration and then pulse centrifuged. The resin was further re-suspended with filtered AAV lysate and incubated by rocking at RT for 10–20 min followed by pulse centrifugation. The resin was then washed thrice with 1 mL of PBS and pulse centrifuged. AAV was eluted twice by resuspending the resin in 1 mL of 0.1 M citric acid (pH 2) (Sigma, Cat#C1909). The supernatant containing AAV particles, obtained through pulse centrifugation, was collected in each elution round.

Purified AAV preps were concentrated using Amicon cassettes (Millipore; Cat#UFC810008). For concentration, the cassettes were first equilibrated with 2 mL of final formulation buffer (FFB) (1,750 μL of 1 M NaCl + 50 μL of 10% pluronic-f68 + 48.2 mL of PBS 1×). Then, purified AAV prep was loaded onto the cassette and centrifuged to concentrate to a volume of 100–200 μL. Three washes of the cassette were performed using 2 mL of FFB. The volume of concentrated AAV prep was measured and was kept in the range of 150–200 μL.

#### AAV production and purification 500-mL scale

HYPERflask-scale AAV preps were produced in Gene Transfer Vector Core at Schepens Eye Research Institute corresponding to data shown in Figures 4B and 4C. HEK293 cells at 80% confluency from four 15-cm dishes were seeded in a HYPERflask (Westnet, Cat#10030) containing complete DMEM. The seeded flasks were incubated at 37°C, 5% $CO_2$ for ∼24 h. After 24 h, pJD64 (Rep2-Cap9), either of pJD34 (AAV genome: CMV-mScarlet) or CMV-eGFP-WPRE, and one of the two helper plasmids (pALD-X80 and pJD171) were mixed in serum-free DMEM to make DNA-DMEM mix. Helpers ALD-X80 and JD171 were used at 1:0.8:1.25 ratio corresponding to Rep2-Cap9:helper:AAV genome for data presented in Figure 4B. Helpers ALD-X80 and JD171 were respectively mixed at 1:0.8:1.25 and 1:1:1.25 ratio corresponding to Rep2-Cap9:helper:AAV genome for data presented in Figure 4C. For example, the mass equivalent for Rep2-Cap9(pJD64):helper(pALD-X80):AAV genome(pJD34) at ratio 1:0.8:1.25 is 260:520:260 μg. A ratio of 1:1.375 was used of DNA (μg) to FectoVIR (μg) for transfection. A mixture of serum-free medium and FectoVIR (Polyplus, Cat#120-100) was prepared and added to the DNA-DMEM mix and allowed to incubate at RT for 15 min. After 15 min, this mixture was added to 500 mL of serum-free medium and used to replace the medium in the HYPERflasks. The flasks were incubated at 37°C with 5% $CO_2$ for 72 h. After 72 h, 5 mL of medium from HYPERflask was replaced with 4.25 mL of DMEM, 0.75 mL of 1 M $MgCl_2$, and 34 μL of benzonase endonuclease (250 U/μL) (Millipore Sigma, Cat#101697) and the HYPERflask was incubated at 37°C for 30 min. Following that, the supernatant from the hyperflask was moved to a 1-L receiver bottle and 34 μL of benzonase endonuclease (250 U/μL) was added to it and incubated at 37°C for 1 h . At the same time, the HYPERflask was treated for 1 h at 37°C with 80 mL of 5M NaCl (Fisher Scientific Cat#BP358-1) to induce lysis. After the incubation, contents from the HYPERflask were transferred to the receiver, and cell debris was pelleted followed by filtration using a 0.45-μm filter (Thermo, Cat#16211-068). The filtered lysate was then purified and concentrated using a Vivaflow200 TFF cassette (Sartorius, Cat#VF20P4) with a molecular weight cut-off (MWCO) of 100 kDa. Following that, the concentrated lysate was loaded into a quick-seal tube, followed by 15%, 25%, 40%, and 55% iodixanol (Serumwerk, Cat#1893) solutions. The tube was ultracentrifuged at 69,000 rpm for 90 min and the 40% and 55% layers were extracted. Final buffer exchange and concentration was achieved using an Amicon Ultra-15 centrifugal filter units (Millipore Sigma, Cat# UFC910008) (MWCO of 100 kDa) to reach a final volume of 1.2 mL.

#### AAV titer determination by qPCR

AAV genome copy number was assessed by qPCR as described earlier[65] using luciferase or minCMV or WPRE primers and TaqMan probes (Table S1). Briefly, crude or purified AAV preps were digested with DNase by mixing following components 5 μL of AAV sample, 5 μL of 10× (1%) Pluronic-F68 (Sigma, Cat#P556), 5 μL of DNase buffer (Roche, Cat#4716728001), 34 μL of Ambion nuclease-free water (Invitrogen, Cat#AM9937), and 1 μL DNase (Roche, Cat#4716728001) and incubated at 37°C for 1 h. DNase-digested samples were diluted 1,000-fold using AAV dilution buffer (1× PCR buffer [Thermo, Cat#4379876], 1× Sheared Salmon Sperm DNA [Thermo, Cat# AM9680], 0.1% Pluronic F-68, Ambion nuclease-free water). qPCR reaction was set up in a 96-well plate (Axygen, Cat#PCR-96-AB-C) by mixing the following components: 2 μL pf primer 1 (JD155 [1 μM] or minCMVfp [9 μM] or WPREfp [9 μM]), 2 μL of primer 2 (JD156 [1.75 μM] or minCMVrp [9 μM] or WPRErp [9 μM]), 2 μL of 2.5-μM TaqMan probe (JD1-104-1 or minCMV or WPRE) 2 μL of Ambion nuclease-free water, 10 μL of 2× TaqMan gene expression master mix (Thermo, Cat#4369016), and 2 μL of diluted sample. The plate was centrifuged at 2,500 × *g* for 30 s. QuantStudio Flex system was used for running the qPCR reaction with the following program: initial denaturation at 95°C for 10 min and then 40 cycles of denaturation at 95°C for 15 s and annealing and extension at 60°C for 60 s. pJD36 (CMV-luciferase) or pJD34 (CMV-mScarlet-WPRE) or CMV-eGFP-WPRE plasmids were linearized using a restriction enzyme and quantified with Qubit. The linearized plasmids were utilized for making the standard curve. Dilutions of the linearized plasmid standards ranged from $10^7$ copies to 10 copies per reaction. Controls for qPCR included validation control (known titer value),

DNase positive control, DNase negative control, water control for DNase digestion, and qPCR water control. GC/mL were calculated as follows:

$$\frac{GC}{mL} = \frac{Quantity * dilution\ factor * Strand\ correction\ factor * mL\ conversion\ factor}{Sample\ volume\ used\ in\ qPCR} \quad \text{(Equation 1)}$$

$$\frac{GC}{mL} = \frac{Quantity * 10000 * 2 * 1000}{2} \quad \text{(Equation 2)}$$

Quantity value was calculated by QuantStudio flex software. An alternative way of getting this value is to plot the $C_t$ values of the standard at different dilutions, get the slope, calculate x for the corresponding y values ($C_t$ values of samples), and raise 10 to the power of those x values giving you quantity. Raw titer values for data shown in different figures can be found in Table S4.

### AAV transduction assay

Transduction efficiency of AAV preps was assessed by measuring luciferase intensity. HEK293 or HeLa cells were seeded respectively at a density of $2 \times 10^4$ and $2.5 \times 10^4$ cells per well in a 96-well plate (Corning, Cat#3904) containing 80 μL of complete DMEM to have around 60% confluency. After 22–24 h, cells in each well were infected with wt-Adenovirus5 (ATCC, Cat#VR-1516) at MOI 20. After 0–2 h, cells were then infected at $10^4$ GC/cell with crude or purified AAV sample. After 48 h, medium was aspirated and 20 μL of 1× lysis buffer (Promega, Cat#E397A) was added to each well (diluted from 5× with $ddH_2O$). The plate was then frozen at −80°C for 60 min and then thawed at 37°C for 15–25 min. The plate was then placed at RT for 5 min. 10 mL of substrate buffer was prepared by mixing the following components: 4.7 mL of $ddH_2O$, 5 mL of Tris-HCl (1 M, pH7.8; Boston Bioproduct, Cat# BBT-78), 30 μL of ATP (100 mM; Thermo, Cat#R0441), $MgCl_2$ (1 M; Sigma, Cat#M1028-100mL or Cat#M8266-1KG), 100× Pierce Firefly Enhancer (Thermo, Cat#16180), and 100 μL of Luciferin (15 mg/mL; Thermo, Cat#88294). Synergy H1 Hybrid plate reader was used for reading luciferase intensity (relative light units [RLUs]) at a gain of 150 after dispensing 100 μL of substrate buffer. RLU values for data shown in different figures can be found in Table S4.

### Capsid ELISA

AAV9 titration ELISA kit (Progen, Cat#PRAAV9) was used as per manufacturer's instructions for determining the quantity of total capsids in the sample (AAV vectors produced in 15-cm dish). Four-parameter logistic curve fit (https://www.aatbio.com/tools/four-parameter-logistic-4pl-curve-regression-online-calculator) was used for calculating capsids/mL (Figure 3D).

### Generation of stable cell lines

HEK293 cells were seeded on a 24-well plate at a density of $6.5 \times 10^4$ cells per well in 500 μL of complete DMEM (supplemented with 10% FBS and 1% penicillin/streptomycin) for achieving a confluency of 60%–70% at the time of transfection. The seeded plates were incubated at 37°C, 5% $CO_2$ for ∼24 h. DNA was transfected using Lipofectamine 2000 (Thermo, Cat#11668-019) at a ratio of 1(μg):3(μL). 300 ng of transfer vector plasmid (pJD179 [*E2A, 22K/33K* and *VA RNA I* helper genes] or pJD180 [*E2A* and *22K/33K* helper genes]) and 100 ng of piggybac transposase plasmid (pTransposase [System biosciences, Cat#P210PA-1]) was transfected for stable integration. Appropriate amounts of plasmids were mixed and the final volume of 50 μL was made using serum-free DMEM. Appropriate volume of Lipofectamine 2000 was taken and serum-free DMEM was used to make final volume of 50 μL (Lipo2000-DMEM mix). The DMEM-DNA mix was incubated for 5 min before mixing with appropriate amount of Lipo2000-DMEM mix. The DNA-Lipo2000 mix was incubated at RT for 10–20 min. 100 μL of DNA-Lipo2000 mix was added dropwise to the respective wells. Medium was replaced with complete DMEM 24 h post transfection. Both transfer vectors possess an inducible mCitrine fluorescent protein. Following integration, cells were expanded either to six-well plate or 10-cm dish and transfected with pJD109 plasmid (Ef1α-ET-pA) in a similar fashion to activate mCitrine expression and enable fluorescence-based sorting of positive integrants.

### Fluorescence-based sorting of positive stable cells

Cells were prepared for sorting in the following manner. Firstly, cells were washed with PBS (1×) and trypsinized by adding 1:1 mix of PBS and trypsin and incubated for 5–10 min at 37°C and 5% $CO_2$. An appropriate amount of complete DMEM ( 3mL for T-75 flask, 100 μL for 24-well plate) was added to neutralize the trypsin. Cells were detached using a pipette and the suspended cell volume was centrifuged at 500 × *g* for 5 min. The supernatant was discarded and the cells re-suspended in PBS-BSA solution (Sigma, Cat#A8806) and kept on ice. The re-suspended cells were filtered through a cell strainer (Corning, Cat#352235) before sorting. The gating strategy for cell sorting can be observed in Figures S5C and S6A. Excitation laser 488 nm and emission filter 530/30 with longpass filter 495 nm and photo multiplier tube (PMT) of 300 mV for mCitrine were used for sorting. Monoclonal and polyclonal cells were recovered in DMEM supplemented with 20% FBS. Cells were expanded and tested for its capacity to produce AAVs.

### Transmission electron microscopy of AAV particles

AAV vectors produced using different helper plasmids in a 15-cm dish and purified via static binding assay were utilized as samples for visualization of viral particles. Firstly, 5 μL of sample was adsorbed onto 200 mesh carbon and formvar-coated nickel grids for 2 min. Three drops of 0.2-μm-filtered distilled water were used to rinse the

grid. Then, the grid was stained by transferring onto a 50-μL drop of 0.2-μm-filtered 2% aqueous uranyl acetate for 30 s. The drop was absorbed off on lens paper and air dried. All grids were imaged using an FEI Tecnai G2 Spirit transmission electron microscope (FEI, Hillsboro, Oregon) at 100-kV accelerating voltage and interfaced with an AMT XR41 digital charge-coupled device (CCD) camera (Advanced Microscopy Techniques, Woburn, MA) for digital TIFF file image acquisition. Transmission electron microscopy images of samples were assessed, and digital images captured at 2k × 2k pixel, 16-bit resolution, 98,000× magnification.

### Flow cytometry

Cells were analyzed on flow cytometer at 48 h post transfection. Transduced HEK293 cells were prepared for flow cytometry by removing the medium and supplying the cells with 1:1 mix of PBS 1×, pH 7.4 (Thermo Fisher, Cat#10010-015) and Accutase (Thermo Fisher, Cat#A11105-01) in a total volume of 100 μL. The cells were incubated for 5–8 min at 37°C, 5% $CO_2$. Cells were then re-suspended and analyzed using BD LSR Fortessa II Cell Analyzer (BD Biosciences). The machine was calibrated with Sphero Rainbow Calibration Particles 8-peak beads (Spherotech, Cat#PCP-30-5A) prior to use. The excitation lasers (Ex) and emission filters (Em) used for respective fluorescent protein measurements were as follows: mScarlet (Ex, 488 nm; Em, 610/20 nm; longpass filter, 595 nm). PMT voltage for fluorescent channel was adjusted in a way that the mean values for 8-peak beads remained constant across different experiments. For reference, PMT for mScarlet was 450 mV.

Bar charts showing the absolute fluorescence units was obtained by analyzing the flow cytometry data using FlowJo software (BD Biosciences). The following steps were undertaken for calculating the absolute units. (1) Live cells were gated based on forward scatter area versus side-scatter area plot. (2) Within the live-cells population, single cells were gated based on forward scatter area versus forward scatter width. (3) Within the single-cells population, mScarlet-positive cells were gated based on a negative control (non-transfected) sample such that 99.9% of cells fell outside of the selected gate. (4) For the mScarlet-positive cell population, FlowJo software calculated mean value of fluorescence and the frequency of positive cells. Multiplying these two values gives absolute intensity, which is a direct measure of the fluorescent protein signal.

$$\textit{Absolute fluorescent intensity} = \textit{Mean fluorescence} * (\%\ \textit{fluorescence positive cells}) \quad \text{(Equation 3)}$$

### RNA extraction

HEK293 cells seeding and transfection in 24-well plates was performed as mentioned earlier. pJD94 (E2A), pJD73 (VA RNA I), pJD143 (E4orf6), DF6, and pJD141 (22K/33K) plasmids were used in transfection. Plasmids were transfected in an equimolar fashion. Junk DNA (pJD03) was used to normalize DNA transfection amount across samples. RNA was extracted from transfected cells after 48 h using Qiagen RNeasy mini kit (Qiagen, Cat#74134) following manufacturer's instructions. After RNA extraction, DNase digestion was performed using Turbo DNA-free kit (Thermo, Cat#AM1907) following manufacturer's instructions. Extracted RNA was accurately quantified using Qbit fluorometer. 1 μg of extracted RNA was reverse transcribed in a total volume of 20 μL using high-capacity cDNA reverse transcription kit (Thermo, Cat#4368814) as per manufacturer guidelines. Random primers provided in the kit were used for reverse transcription. Then, qPCR was performed as described earlier. 2 μL of cDNA sample was used for qPCR. Human 18S rRNA (h18SrRNA) was used as control gene. Delta delta $C_t$ method was used for calculating relative mRNA expression. Control conditions included water as template both for reverse transcription and qPCR reactions as well as samples with no reverse transcription. Refer to Table S1 regarding sequences and desired concentrations of TaqMan probes and primers used in qPCR for *H18SrRNA*, *E2A*, *E4orf6*, and *VA RNA I* genes.

### Copy-number analysis using ddPCR

Genomic DNA was extracted from HEK293 cells and 2M4 monoclonal cells using DNeasy Blood and Tissue kit (Qiagen, Cat#69506). Genomic DNA was subsequently digested with BamHI restriction enzyme and purified using Qiaquick purification kit (Qiagen, Cat#28104). The purified genomic DNA samples were then quantified using Qubit. 5 ng of genomic DNA was used as template for ddPCR. E2A primers and probe were utilized for ddPCR (Table S1). No-template control condition used $ddH_2O$ as template. ddPCR reaction was performed as mentioned in Sanmiguel et al.[65]

### Statistical analysis

All experimental conditions were repeated at least twice. Biological replicates ($n$ = 2 or 3) were measured for most of the samples/conditions. Exceptions include production runs at 18-mL scale (Figures 3C and 3D) and 500-mL scale (Figures 4B–4D) that utilized single biological replicate. qPCR or capsid ELISA or transduction assay of AAV prep samples was either done in duplicates or triplicates. All error bars represent mean ± standard deviation. Figure legends have relevant details. MS Excel and Graphpad Prism software were used for analyzing data, making charts, and performing statistical analysis. One-way ANOVA with Tukey's *post hoc* test and Wilcoxon signed rank test were utilized for statistical analysis. Details regarding data subjected to statistical analysis can be found in the figure legends.

## DATA AND CODE AVAILABILITY

Some plasmids genereated in this study have been deposited on Addgene and are available for non-commercial academic use. Other plasmids and cell-line reagents generated in this study can be made available upon request to the corresponding author.

## ACKNOWLEDGMENTS

Part of this work was funded by a Sponsored Research Agreement with Akouos, Inc. We would like to thank MGH HSCI flow cytometry core, SERI flow cytometry core, MGH DNA sequencing core, Philip Seiffert from SERI Morphology core, and Timothy Vanderleest for assistance with experiments and analysis. We would like to thank Matthew Weitzman for feedback on the manuscript. We would also like to thank Achla Gandhi, Ainsley Maclachlan, Julio Sanmiguel, and other members of the Zabaleta lab for assistance in conducting experiments and for meaningful discussions.

## AUTHOR CONTRIBUTIONS

J.D. conceived the research project, performed and designed most of the experiments, analyzed data, and wrote the paper. L.H.V. conceived and supervised the research project and wrote the paper. N.Z. designed experiments, supervised the research project, and wrote the paper. J.S. performed cloning and a few early experiments. E.C. performed production experiments at HYPERflask scale.

## DECLARATION OF INTERESTS

J.D., L.H.V., and N.Z. are inventors in a patent that protects these findings. N.Z. is the PI in a sponsored research agreement with Akouos, Inc., that partially funded this work. L.H.V. is co-founder, board member, and equity holder to Affinia Therapeutics, an AAV gene-therapy company. He further is an advisor to Eli Lilly and company, Board member to Odylia and Addgene, and co-inventor on various AAV and gene-therapy-related patents.

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2024.101376.

## REFERENCES


1. Issa, S.S., Shaimardanova, A.A., Solovyeva, V.V., and Rizvanov, A.A. (2023). Various AAV Serotypes and Their Applications in Gene Therapy: An Overview. Cells *12*, 785.
2. Barrett, D., Foss-Campbell, B., Wendland, A., Nguyen-Jatkoe, L., Micklus, A., and Shin, D. (2021). Gene, Cell, & RNA therapy landscape Q2 2021 quarterly data report (ASGCT), p. 21.
3. Sha, S., Maloney, A.J., Katsikis, G., Nguyen, T.N.T., Neufeld, C., Wolfrum, J., Barone, P.W., Springs, S.L., Manalis, S.R., Sinskey, A.J., and Braatz, R.D. (2021). Cellular pathways of recombinant adeno-associated virus production for gene therapy. Biotechnol. Adv. *49*, 107764.
4. Rininger, J., Fennell, A., Schoukroun-Barnes, L., Peterson, C., and Speidel, J. (2019). Capacity analysis for viral vector manufacturing. Bioproc. Int. *17*, 2–9.
5. Louis, N., Evelegh, 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.
6. Lin, Y.C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., et al. (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. Nat. Commun. *5*, 4767.
7. Aponte-Ubillus, J.J., Barajas, D., Peltier, J., Bardliving, C., Shamlou, P., and Gold, D. (2018). Molecular design for recombinant adeno-associated virus (rAAV) vector production. Appl. Microbiol. Biotechnol. *102*, 1045–1054.
8. E Cameau, C.G. (2020). A Pedregal Cost modelling comparison of adherent multi-trays with suspension and fixed-bed bioreactors for the manufacturing of gene therapy products. Cell Gene Ther. Insights *5*, 1663–1674.
9. Giles, A., Lock, M., Chen, S.J., Turner, K., Wesolowski, G., Prongay, A., Petkov, B.N., Olagbegi, K., Yan, H., and Wilson, J.M. (2023). Significant Differences in Capsid Properties and Potency Between Adeno-Associated Virus Vectors Produced in Sf9 and HEK293 Cells. Hum. Gene Ther. *34*, 1003–1021.
10. Wright, J.F. (2023). AAV vector production: Troublesome host innate responses in another setting. Mol. Ther. Methods Clin. Dev. *28*, 412–413.
11. 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.
12. Matsushita, T., Elliger, S., Elliger, C., Podsakoff, G., Villarreal, L., Kurtzman, G.J., Iwaki, Y., and Colosi, P. (1998). Adeno-associated virus vectors can be efficiently produced without helper virus. Gene Ther. *5*, 938–945.
13. Adsero, A., Chestnut, B., Shahnejat-Bushehri, S., Sasnoor, L., McMurphy, T., Swenor, M., Pasquino, R., Pradhan, A., Hernandez, V., Padegimas, L., and Dismuke, D. (2024). A Novel Role for the Adenovirus L4 Region 22K and 33K Proteins in Adeno-Associated Virus Production. Hum. Gene Ther. *35*, 59–69.
14. Johari, Y.B., Pohle, T.H., Whitehead, J., Scarrott, J.M., Liu, P., Mayer, A., and James, D.C. (2024). Molecular design of controllable recombinant adeno-associated virus (AAV) expression systems for enhanced vector production. Biotechnol. J. *19*, e2300685.
15. Pilder, S., Moore, M., Logan, J., and Shenk, T. (1986). The adenovirus E1B-55K transforming polypeptide modulates transport or cytoplasmic stabilization of viral and host cell mRNAs. Mol. Cell Biol. *6*, 470–476.
16. Woo, J.L., and Berk, A.J. (2007). Adenovirus ubiquitin-protein ligase stimulates viral late mRNA nuclear export. J. Virol. *81*, 575–587.
17. Samulski, R.J., and Shenk, T. (1988). Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. J. Virol. *62*, 206–210.
18. Janik, J.E., Huston, M.M., Cho, K., and Rose, J.A. (1989). Efficient synthesis of adeno-associated virus structural proteins requires both adenovirus DNA binding protein and VA I RNA. Virology *168*, 320–329.
19. Mathews, M.B., and Shenk, T. (1991). Adenovirus virus-associated RNA and translation control. J. Virol. *65*, 5657–5662.
20. Wahid, A.M., Coventry, V.K., and Conn, G.L. (2008). Systematic deletion of the adenovirus-associated RNAI terminal stem reveals a surprisingly active RNA inhibitor of double-stranded RNA-activated protein kinase. J. Biol. Chem. *283*, 17485–17493.
21. Machitani, M., Yamaguchi, T., Shimizu, K., Sakurai, F., Katayama, K., Kawabata, K., and Mizuguchi, H. (2011). Adenovirus Vector-Derived VA-RNA-Mediated Innate Immune Responses. Pharmaceutics *3*, 338–353.
22. Chung, C.H., Murphy, C.M., Wingate, V.P., Pavlicek, J.W., Nakashima, R., Wei, W., McCarty, D., Rabinowitz, J., and Barton, E. (2023). Production of rAAV by plasmid transfection induces antiviral and inflammatory responses in suspension HEK293 cells. Mol. Ther. Methods Clin. Dev. *28*, 272–283.
23. 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.
24. Aksu Kuz, C., Ning, K., Hao, S., Cheng, F., and Qiu, J. (2024). Role of the membrane-associated accessory protein (MAAP) in adeno-associated virus (AAV) infection. J. Virol. *98*, e0063324.
25. Elmore, Z.C., Patrick Havlik, L., Oh, D.K., Anderson, L., Daaboul, G., Devlin, G.W., Vincent, H.A., and Asokan, A. (2021). The membrane associated accessory protein is an adeno-associated viral egress factor. Nat. Commun. *12*, 6239.
26. Swaminathan, S., and Thimmapaya, B. (1996). Transactivation of adenovirus E2-early promoter by E1A and E4 6/7 in the context of viral chromosome. J. Mol. Biol. *258*, 736–746.
27. Gilardi, P., and Perricaudet, M. (1986). The E4 promoter of adenovirus type 2 contains an E1A dependent cis-acting element. Nucleic Acids Res. *14*, 9035–9049.
28. Chang, L.S., Shi, Y., and Shenk, T. (1989). Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. J. Virol. *63*, 3479–3488.
29. Stracker, T.H., Cassell, G.D., Ward, P., Loo, Y.M., van Breukelen, B., Carrington-Lawrence, S.D., Hamatake, R.K., van der Vliet, P.C., Weller, S.K., Melendy, T., and Weitzman, M.D. (2004). The Rep protein of adeno-associated virus type 2 interacts with single-stranded DNA-binding proteins that enhance viral replication. J. Virol. *78*, 441–453.
30. Ward, P., Dean, F.B., O'Donnell, M.E., and Berns, K.I. (1998). Role of the adenovirus DNA-binding protein in *in vitro* adeno-associated virus DNA replication. J. Virol. *72*, 420–427.
31. Cathomen, T., and Weitzman, M.D. (2000). A functional complex of adenovirus proteins E1B-55kDa and E4orf6 is necessary to modulate the expression level of p53 but not its transcriptional activity. J. Virol. *74*, 11407–11412.
32. Liu, Y., Shevchenko, A., Shevchenko, A., and Berk, A.J. (2005). Adenovirus exploits the cellular aggresome response to accelerate inactivation of the MRN complex. J. Virol. *79*, 14004–14016.
33. Nevels, M., Rubenwolf, S., Spruss, T., Wolf, H., and Dobner, T. (2000). Two distinct activities contribute to the oncogenic potential of the adenovirus type 5 E4orf6 protein. J. Virol. *74*, 5168–5181.
34. Querido, E., Blanchette, P., Yan, Q., Kamura, T., Morrison, M., Boivin, D., Kaelin, W.G., Conaway, R.C., Conaway, J.W., and Branton, P.E. (2001). Degradation of

p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. Genes Dev. *15*, 3104–3117.

35. 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.
36. 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.
37. Babiss, L.E., and Ginsberg, H.S. (1984). Adenovirus type 5 early region 1b gene product is required for efficient shutoff of host protein synthesis. J. Virol. *50*, 202–212.
38. Babiss, L.E., Ginsberg, H.S., and Darnell, J.E., Jr. (1985). Adenovirus E1B proteins are required for accumulation of late viral mRNA and for effects on cellular mRNA translation and transport. Mol. Cell Biol. *5*, 2552–2558.
39. Fisher, K.J., Gao, G.P., Weitzman, M.D., DeMatteo, R., Burda, J.F., and Wilson, J.M. (1996). Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. J. Virol. *70*, 520–532.
40. Nayak, R., and Pintel, D.J. (2007). Positive and negative effects of adenovirus type 5 helper functions on adeno-associated virus type 5 (AAV5) protein accumulation govern AAV5 virus production. J. Virol. *81*, 2205–2212.
41. Lu, S., and Cullen, B.R. (2004). Adenovirus VA1 noncoding RNA can inhibit small interfering RNA and MicroRNA biogenesis. J. Virol. *78*, 12868–12876.
42. Aparicio, O., Carnero, E., Abad, X., Razquin, N., Guruceaga, E., Segura, V., and Fortes, P. (2010). Adenovirus VA RNA-derived miRNAs target cellular genes involved in cell growth, gene expression and DNA repair. Nucleic Acids Res. *38*, 750–763.
43. Aparicio, O., Razquin, N., Zaratiegui, M., Narvaiza, I., and Fortes, P. (2006). Adenovirus virus-associated RNA is processed to functional interfering RNAs involved in virus production. J. Virol. *80*, 1376–1384.
44. Andersson, M.G., Haasnoot, P.C.J., Xu, N., Berenjian, S., Berkhout, B., and Akusjärvi, G. (2005). Suppression of RNA interference by adenovirus virus-associated RNA. J. Virol. *79*, 9556–9565.
45. Su, W., Seymour, L.W., and Cawood, R. (2023). AAV production in stable packaging cells requires expression of adenovirus 22/33K protein to allow episomal amplification of integrated rep/cap genes. Sci. Rep. *13*, 21670.
46. Wu, K., Orozco, D., and Hearing, P. (2012). The adenovirus L4-22K protein is multifunctional and is an integral component of crucial aspects of infection. J. Virol. *86*, 10474–10483.
47. Morris, S.J., and Leppard, K.N. (2009). Adenovirus serotype 5 L4-22K and L4-33K proteins have distinct functions in regulating late gene expression. J. Virol. *83*, 3049–3058.
48. Wu, K., Guimet, D., and Hearing, P. (2013). The adenovirus L4-33K protein regulates both late gene expression patterns and viral DNA packaging. J. Virol. *87*, 6739–6747.
49. Lan, S., Kamel, W., Punga, T., and Akusjärvi, G. (2017). The adenovirus L4-22K protein regulates transcription and RNA splicing via a sequence-specific single-stranded RNA binding. Nucleic Acids Res. *45*, 1731–1742.
50. Nie, Y., Pan, H., Li, Q., Na, H., Figueroa, B., and Vincent, K. (2024). Characterization of the function of Adenovirus L4 gene products and their impact on AAV vector production. Mol. Ther. Methods Clin. Dev. *32*, 101370. https://doi.org/10.1016/j.omtm.2024.101370.
51. White, L., Erbay, B., and Blair, G.E. (2023). The Cajal body protein p80-coilin forms a complex with the adenovirus L4-22K protein and facilitates the nuclear export of adenovirus mRNA. mBio *14*, e0145923.
52. Qiao, C., Li, J., Skold, A., Zhang, X., and Xiao, X. (2002). Feasibility of generating adeno-associated virus packaging cell lines containing inducible adenovirus helper genes. J. Virol. *76*, 1904–1913.
53. Lee, Z., Lu, M., Irfanullah, E., Soukup, M., and Hu, W.S. (2022). Construction of an rAAV Producer Cell Line through Synthetic Biology. ACS Synth. Biol. *11*, 3285–3295.
54. 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.
55. Kruger-Haag, A., Lehmann, C., Schmidt, E., Sonntag, F., Horer, M., and Kochanek, S. (2019). Evaluation of life cycle defective adenovirus mutants for production of adeno-associated virus vectors. J. Gene Med. *21*, e3094.
56. Zabaleta, N., Bhatt, U., Hérate, C., Maisonnasse, P., Sanmiguel, J., Diop, C., Castore, S., Estelien, R., Li, D., Dereuddre-Bosquet, N., et al. (2022). Durable immunogenicity, adaptation to emerging variants, and low-dose efficacy of an AAV-based COVID-19 vaccine platform in macaques. Mol. Ther. *30*, 2952–2967.
57. Weber, W., Fux, C., Daoud-el Baba, M., Keller, B., Weber, C.C., Kramer, B.P., Heinzen, C., Aubel, D., Bailey, J.E., and Fussenegger, M. (2002). Macrolide-based transgene control in mammalian cells and mice. Nat. Biotechnol. *20*, 901–907.
58. Guerin, K., Rego, M., Bourges, D., Ersing, I., Haery, L., Harten DeMaio, K., Sanders, E., Tasissa, M., Kostman, M., Tillgren, M., et al. (2020). A Novel Next-Generation Sequencing and Analysis Platform to Assess the Identity of Recombinant Adeno-Associated Viral Preparations from Viral DNA Extracts. Hum. Gene Ther. *31*, 664–678.
59. Tai, P.W.L., Xie, J., Fong, K., Seetin, M., Heiner, C., Su, Q., Weiand, M., Wilmot, D., Zapp, M.L., and Gao, G. (2018). Adeno-associated Virus Genome Population Sequencing Achieves Full Vector Genome Resolution and Reveals Human-Vector Chimeras. Mol. Ther. Methods Clin. Dev. *9*, 130–141.
60. Liu, S., Li, J., Peraramelli, S., Luo, N., Chen, A., Dai, M., Liu, F., Yu, Y., Leib, R.D., Li, Y., et al. (2024). Systematic comparison of rAAV vectors manufactured using large-scale suspension cultures of Sf9 and HEK293 cells. Mol. Ther. *32*, 74–83.
61. Ferrari, F.K., Samulski, T., Shenk, T., and Samulski, R.J. (1996). Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. J. Virol. *70*, 3227–3234.
62. Schwartz, R.A., Palacios, J.A., Cassell, G.D., Adam, S., Giacca, M., and Weitzman, M.D. (2007). The Mre11/Rad50/Nbs1 complex limits adeno-associated virus transduction and replication. J. Virol. *81*, 12936–12945.
63. Zhou, H., O'Neal, W., Morral, N., and Beaudet, A.L. (1996). Development of a complementing cell line and a system for construction of adenovirus vectors with E1 and E2a deleted. J. Virol. *70*, 7030–7038.
64. Florea, M., Nicolaou, F., Pacouret, S., Zinn, E.M., Sanmiguel, J., Andres-Mateos, E., Unzu, C., Wagers, A.J., and Vandenberghe, L.H. (2023). High-efficiency purification of divergent AAV serotypes using AAVX affinity chromatography. Mol. Ther. Methods Clin. Dev. *28*, 146–159.
65. Sanmiguel, J., Gao, G., and Vandenberghe, L.H. (2019). Quantitative and Digital Droplet-Based AAV Genome Titration. Methods Mol. Biol. *1950*, 51–83.