



Generation and characterization of a HEK293 cell line optimized for recombinant adeno-associated virus production

Mélissa Vona¹, Iris Bodenmann¹, Marc-Antoine Perrenoud, Rachel Buchs, Pelin Kolcak Yasli, Luca Nanni, Romain Daveau, Alexandre Félix, Jens Stolte, Ann-Kristin Hov, Bertrand Chollet, Thierry Schuepbach, Déborah Ley, Efrain Guzman^{*}, Igor Fisch, Nicolas Mermoud

NewBiologix SA, Route de la Corniche 6-8, 1066 Epalinges, Switzerland

ARTICLE INFO

Keywords:

Cell line development
Viral vector production
Adeno-associated viral vectors
Human cell lines

ABSTRACT

HEK293 is a preferred cellular platform to produce viral vectors including adeno-associated viruses (AAV). However, HEK293 cells were shown to be genomically unstable and many HEK293 cell lines having distinct genotypes and phenotypes have been reported. Here we generated a stable clonal cell line specifically selected for the optimal production of recombinant AAV (rAAV) by the triple plasmid transfection method. Initially over two thousand single cell clones were isolated from a HEK293 polyclonal cell line and evaluated for their growth profile in suspension, doubling time, ability to recover freeze-thaw cycles and transfection efficacy. A selection of clones that met these specific criteria were then screened for their ability to produce high rAAV titers by triple plasmid transfection, yielding one high-performing clone named NBX1P01. This clone was genomically characterized using optical genome mapping and whole genome sequencing and further evaluated for rAAV production capacity across different serotypes and genes of interest (GOI). NBX1P01 was shown to be genomically stable over 55 population doubling levels (PDL), highly transfectable and able to produce rAAV titers similar or higher than those produced by a commercially available HEK293 cell line using the same culture, transfection, harvest and quantification protocol. The ratio of full-to-empty rAAV particles produced by NBX1P01 was two-fold higher than those of the commercial cell line. Long-read sequencing of the encapsidated DNA from the NBX1P01-produced rAAV indicated high levels of genome integrity with minimal levels of contaminants. These results demonstrated the versatility of NBX1P01 cells and their ability to produce high-quality rAAV vectors.

1. Introduction

The HEK293 cell line was established by transforming human embryonic kidney cells with sheared adenovirus type 5 DNA [1], and many subtypes, subclones and derivatives have been generated. The parental HEK293 and its -T, and -F derivatives are frequently used in the production of biopharmaceuticals [2]. Because of their ease of culture, manipulation and transfectability, HEK293 were used for the transient expression of recombinant proteins. For instance, NUWIQ®, a recombinant coagulation factor VIII (FVIII), is among the most commonly known therapeutics produced in HEK293F cells [3,4].

HEK293 is the preferred cellular platform for the production of recombinant adeno-associated virus (rAAV) viral vectors. This stems from the presence of genomically integrated adenoviral E1A/B genes which

provide helper functions during viral vector production [5]. AAV are non-enveloped, single-stranded DNA viruses which have shown safety and efficacy as gene therapy recombinant viral vectors [6]. rAAVs are typically produced by the transient transfection of HEK293 cells with three plasmids containing: 1) AAV replicase (*rep*) and capsid (*cap*) genes, 2) adenovirus E4, E2, and VA helper genes, and 3) the transgene of interest flanked by AAV inverted terminal repeat (ITR) sequences.

HEK293 cells are known to be genomically unstable [7]. Thus, although all the different HEK293 derivatives available to date were derived from the same original cell line, passaging over many years, changes in culture conditions, adaptation to high-density growth in serum-free conditions and its genomic instability resulted in the generation of diverse pools of HEK293 cells with distinct genotypes and phenotypes [8]. Moreover, these HEK293 derivatives are not necessarily

^{*} Corresponding author.

E-mail address: efrain.guzman@newbiologix.com (E. Guzman).

¹ Both authors contributed equally to this work

optimal for rAAV production across all serotypes and transgenes, as their productivity, transfectability, and rAAV yield can vary significantly. Here we describe the generation of a clonal cell line specifically selected for the optimal production of rAAV by the triple plasmid transient transfection method [9]. Over two thousand single cell clones were derived from an original HEK293 polyclonal cell line and evaluated for their ability to be transfected and optimal growth profiles. A selection of these clones was then screened for their ability to produce high titers of rAAV and high full-to-empty capsid ratios by the triple plasmid transfection method. One high-performing and one low-performing clones were then selected and evaluated for rAAV production capacity in small and medium scale cultures. These clones were also genomically characterized and their stability was assessed upon further culturing.

2. Materials and methods

2.1. Cells and media

Suspension-adapted HEK293 cells were obtained from Florabio AS (Turkey) and grown in ORCHID S media (Florabio) supplemented with 2 mM L-glutamine (Gibco) or Viral Production Media supplemented with 4 mM GlutaMAX (Thermo Fisher Scientific) and anticlumping agent (Gibco). Cells were maintained in a volume of 20 mL in 50 mL TubeSpin Bioreactors (TPP) at 37°C, 225 rpm (50 mm rotor) under 7 % CO₂, 85 % humidity in a shaker-incubator (Kühner) and were subcultured every 3–4 days. Cell viabilities and viable cell densities (VCD) were quantified using a Vi-CELL BLU (Beckman). To generate single-cell clones, cells were dispensed and imaged using an UP.SIGHT (Cytena) following the manufacturer's instructions, and the growth of the subclones was monitored using an automated microscopic imaging system, the Cella-vista 4 (SYNTEC). Population doubling level (PDL) zero (PDL 0) was defined as the time at which the stability study started. For comparison studies, Expi293F and Viral Production Cells 2.0 (VPC2.0, Thermo Fisher Scientific) were cultured following the manufacturer's instructions.

2.2. Recombinant AAV vector production, purification and titration

Rep2/Cap1/2/5/8, Helper and Genome plasmids were purchased from Takara-Bio. rAAV Genome plasmid contained an enhanced green fluorescent protein (EGFP) under the control of a CMV promoter. Other promoter-transgene combinations were synthesized by VectorBuilder (Table 1). Small-scale rAAV productions were performed by the triple transfection method [9,10] under agitation conditions in shallow-well 24-well plates (EnzyScreen). Medium-scale vector productions were performed in TubeSpin Bioreactor 50 (TPP) and 1.5 L productions in disposable 3L-Erlenmeyer culture flasks (TriForest). Briefly, cells were subcultured at 3×10^5 cells/ml and grown for 3 or 4 days under defined incubation conditions. The day before transfections, cells were seeded in the corresponding vessels at 5×10^5 cells/ml. Plasmid DNA (molar ratio

of 1:2:2 of Helper:Rep2/CapX:Genome) and transfection reagent (Polyethylenimine, Linear, MW 25,000 (PEI 25 K), from Polysciences) were mixed at a ratio of 1:2 DNA:Reagent, and added to the cells after a complexation time of 5 min. Transiently transfected cells were cultured for 72 hrs at 37°C, 225 rpm under 7 % CO₂, 85 % humidity.

2.3. Purification of rAAV particles

The HEK293 cells were lysed post-cultivation using a 1 % polysorbate 20-based lysis buffer containing 3 U/mL Salt-Activated Nuclease (SAN, ArcticZymes Technologies). Following a 2-hour incubation at 37°C with agitation, the lysates were centrifuged at 700 rpm for 25 min at 4°C, allowing cell debris to precipitate. The supernatant containing rAAV particles was carefully separated, filtered through a 0.22 µm filter, and loaded on a POROS GoPure™ AAVX Pre-packed affinity column. Column washing proceeded with 10 vol of 20 mM Tris, 1.5 mM NaCl (pH 7.5) until the absorbance at 280 nm reached baseline, using an ÄKTA go HPLC system (Cytiva). The bound rAAV vectors were eluted with 100 mM glycine, 100 mM NaCl (pH 2.5) and immediately neutralized with 1/20 vol of 1 M Tris-Base (pH 9.0). For further enrichment of full rAAV particles, ion exchange chromatography was performed using a 1 mL POROS™ GoPure™ XQ column (Thermo Fisher Scientific), equilibrated with 20 mM Tris, 5 mM MgSO₄ (pH 9.0), and eluted in a 2-step gradient with Buffer B (50 mM Tris, 5 mM MgSO₄, pH 9.0) at 22 % and 100 % over 20 column volumes.

2.4. Quantification of viral vector genome titers by dPCR

Cell suspensions were freeze-thawed three times and then centrifuged at 3000 x g for 10 min at room temperature to pellet cell debris. Supernatants were harvested as crude rAAV preparations. Extra-viral DNA was digested with DNase I (Ambion) following the manufacturer's instructions. After DNase I inactivation, rAAV genomes were isolated using the CGT Viral Vector Lysis kit (Qiagen) following the manufacturer's instructions. The viral vector DNA was diluted and mixed with a dPCR Master Mix (QIAcuity Probe PCR kit - Qiagen) containing a Cy5-labeled Primer Probe mix targeting the CMV promoter (forward primer: 5'-CCCAGTACATGACCTTATGGGA-3', reverse primer: 5'-GCCCATTGATGTACTGCCAAA-3', Taqman probe: 5'-TCGCTATTAC-CATGGTGATGCGGT-3') and transferred to a QIAcuity Nanoplate (Qiagen). Primers and probes used for titrating transgenes other than EGFP (Table 1) were designed using Primer3 and synthesized by MicroSynth AG. Viral vector genomes were amplified by using a QIAcuity One instrument (Qiagen) and the data analyzed using QIAcuity software Suite version 2.1.7.182. Titers are reported as vector genomes per milliliter (vg/ml).

2.5. Characterization of rAAV full vs empty capsid ratio and genome integrity

Mass photometry analyses of empty and DNA-containing capsids were carried out using a SamuxMP (Refeyn) as described previously [11, 12] and following the manufacturer's instructions.

Recombinant AAV genome quality was assessed using Oxford Nanopore Technologies (ONT) long-read sequencing as described in the Supplementary Information. Bioinformatic analysis of the sequenced reads was performed as described previously [13].

2.6. rAAV genome analysis by long-read sequencing

Recombinant AAV genome quality was assessed using Oxford Nanopore Technologies (ONT) long-read sequencing. Purified rAAV capsids were first assessed for genomic and plasmid DNA contamination using the Qubit ssDNA and dsDNA assay (Thermo Fisher Scientific) following the manufacturer's instructions. Encapsidated rAAV genomes were extracted using the PureLink™ Viral RNA/DNA Mini Kit (Thermo

Table 1
Transgene cassettes used for rAAV vector construction and production.

Promoter	Transgene	Reference	PolyA	Size, ITR-to-ITR (bp)
CBA	mRPE65	NM_029987.2	BGH	3038
TBG	BDD hFVIII	KY682701.1	BGH	5420
CMV	Aflibercept	DB08885	BGH	2474
CBA	LacZ	KC896840.1	BGH	4487
SPC5-12	Nano Luciferase	JQ437370.1	BGH	1447
CMV+intron	EGFP	MH325108.1	BGH	2811

CBA: Chicken beta-actin; TBG: thyroxine-binding globulin; CMV: cytomegalovirus; SPC5-12: synthetic muscle-specific; mRPE65: mouse Retinal pigment epithelium-specific 65 kDa protein; BDD hFVIII: B-domain-deleted human clotting factor VIII; EGFP: enhanced green fluorescent protein; BGH: bovine growth hormone.

Fisher Scientific) following manufacturer's instructions. The extracted DNA was analyzed for quantity and quality using Qubit and FemtoPulse assays to confirm DNA concentration and expected DNA fragment profile.

Library preparation was performed according to the Ligation sequencing amplicons - Native Barcoding Kit 24 V14 protocol from Oxford Nanopore. Samples were barcoded using the Native Barcoding Kit 24 V14 (Oxford Nanopore) and pooled together. Sequencing was performed on a MinIon Flowcell following standard procedure and run recommendations defined by the manufacturer for sequencing. Bioinformatic analysis of the sequenced reads was performed as described previously [13].

2.7. Genotypic and genomic characterization of cell clone stability

Optical genome mapping (OGM) was carried out using Bionano Genomics Saphyr system. Single-molecule imaging raw data were processed using Access software version 1.8.2 and hg38.DLE1.0k-b_0labels_masked_YPARs.cmap file as the reference genome. First, consistency across samples of molecule quality metrics such as Direct Labeling Enzyme 1 (DLE1) density (14–16/100 Kb), N50 (230–330 Kb), total number of aligned molecules (4.4–5.3 M) and effective coverage (260–400X) of the reference genome were checked. Subsequently, the Rare Variant Analysis pipeline was launched on each sample with neither copy-number variants (CNV) nor structural variations (SV) mask filtering applied from the Advanced Analysis Option panel. Custom R scripts were used to perform data post-processing and analysis for chromosome-scaled ploidy estimates chromosome-specific coverage and variances, translocations, duplications, inversions and insertions/deletions. The genomic stability of a given clone over time was assessed by comparing changes in called variants between the parental HEK293 cell line and the clone at generation 0 on the one hand, and the clone between generations 0 and 55 on the other. For structural variant analysis, a Fisher's exact test was used to evaluate significant discrepancies in variant counts between the parental cell line and its derived clones. For copy number variant analysis, all non-overlapping genomic intervals interested in copy number changes were tested for significant differences of the impacted total genome size using a Pearson's Chi-squared test with continuity correction.

2.8. Clonality assessment, single-nucleotide polymorphism and structural variant calling from PacBio sequencing data

Whole-genome sequencing was performed following PacBio recommendations. Briefly, high-molecular weight DNA was extracted using PacBio Nanobind CBB extraction kit (Pacific Biosciences). DNA shearing was performed with Megaruptor 3, producing on average 15–18 Kb-long fragments. Library preparation was performed according to the SMRTbell prep kit 3.0 (PacBio) following manufacturer's instructions. Sequencing was performed on a Revio instrument (Pacific Biosciences) following standard procedures defined by the manufacturer following the SMRTLink sample setup and run recommendations.

HiFi reads were aligned to the Telomere-to-Telomere (T2T) reference genome [14] using minimap2 [15]. Single-nucleotide polymorphism (SNPs) calling was performed with DeepVariant [16] with T2T as control reference for each cell line. Multi-allelic variants, and variants overlapping with repeated or low-complexity regions were detected using RepeatMasker (<http://www.repeatmasker.org>), and variants that were detected in the parental polyclonal population with an allele frequency of 1, indicating differences with the T2T reference only related to assembly mismatch, were excluded from further analysis. Structural variations (SVs) were detected using pbsv (Pacific Biosciences) and underwent the same screening as SNPs. Clonality assessment was performed as previously described [17], as based on a putatively haploid region on the polyclonal HEK293 cell line identified from copy number analysis (chr15:20Mb-30Mb) and the distribution of Variant Allele

Frequencies (VAF) scores for the SNPs along that region. Clonal populations showed an enrichment in SNPs with VAF close to 1. Copy number analysis on PacBio reads was performed by computing the read coverage in 1Mb bins along the T2T genome. Normalized read coverage for each bin was computed dividing the raw coverage by the median genomic coverage and applying Log₁₀ normalization.

Whole-genome data obtained from PacBio sequencing was used to assemble the genome of cell clones. Specifically, PacBio's HiFi long-reads were assembled de-novo using Hifiasm [18]. Polishing and chromosome-scaled scaffolding of the raw assembly were performed using RagTag [19]. Genome-wide copy number changes, low-complexity (non-)syntenic regions, (un-)mapped genes, transcript and (un-)methylated CpGs were respectively called using deepTools [20], DNACopy (<https://github.com/veseshan/DNACopy>), nQuire [21], windowmasker [22], syri [23], Liftoff [24] and MethBat (<https://github.com/PacificBiosciences/MethBat>) respectively. Sequencing and assembly QA/QC metrics were evaluated using fastp [25], merquy [26] and BUSCO [27] software tools.

2.9. Statistics

Geographic statistics such as the means, standard deviations (S.D.), standard errors of the means (S.E.M) and significance (p) were graphed and calculated using GraphPad Prism for Windows. Unless indicated, significant differences were calculated using 2-way analysis of variants (ANOVA).

3. Results

3.1. Generation of a clonal cell line with an enhanced rAAV production profile

The selection process designed to screen and identify HEK293 cell subclones displaying optimized rAAV production and cell culture properties is shown in Fig. 1. To generate single cell clones from the parental HEK293 cell population, 2298 single cells were printed onto 384-well plates and expanded sequentially up to 50 mL TubeSpin suspension cultures. Ninety-nine single cell clones survived the process and a 2-cycle evaluation process for growth profiles, transfectability using both chemical transfection and electroporation, and survival ability to freeze-thawing. Cell clones displaying high transfection efficacies, yielding a high proportion of the cells expressing a chemically transfected transgene were identified and selected for further analysis (Supplementary Fig. S1).

The ability of selected cell clones to produce high rAAV titers by triple plasmid transfection were investigated, resulting in the selection of one highly performing (HP) clone displaying improved properties and one low performance (LP) clone used as a control. The performance of the HP clone, named NBX1P01, and of the LP control clone NBX6M16, were compared to the parental HEK293 cells. NBX1P01 showed similar transfection efficiency as the original polyclonal HEK293 cells but significantly higher transfection efficiency when compared to NBX6M16 (Fig. 2A). The doubling time (DT) of NBX1P01 increased from 26 h to 29 h over 55 population doubling level (PDL); in contrast, the DT for both the parental HEK293 and clone NBX6M16 decreased from 30 h to 27 h and 27 h to 26 h respectively (Fig. 2B). The viability of all cells was >95 % over 55 PDL (data not shown). During initial rAAV production studies, the HP clone produced more rAAV2 ($p < 0.0001$) than the parental HEK293 cells. In contrast, the LP clone produced significantly less rAAV1, 2, 5 and 8 than the HP clone NBX1P01 (Fig. 2C). To demonstrate the performance of NBX1P01 at various cell suspension culture scales, rAAV8 was generated in increasing culture volumes, from 10 mL up to 1.5 L and in all cases vector production was consistently similar (Fig. 2D). These studies demonstrated the scalability and versatility of NBX1P01 for rAAV production.

To demonstrate the capacity of NBX1P01 to produce rAAV of various

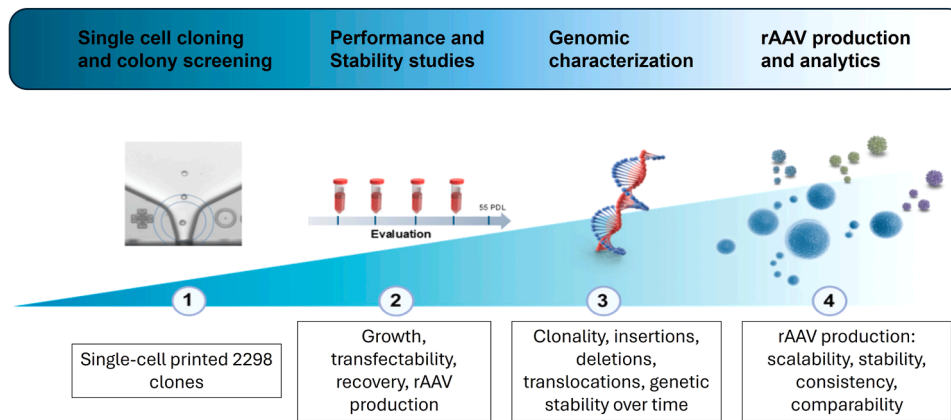


Fig. 1. Schematic representation of the selection process carried out for the selection of high-performing HEK293 cells.

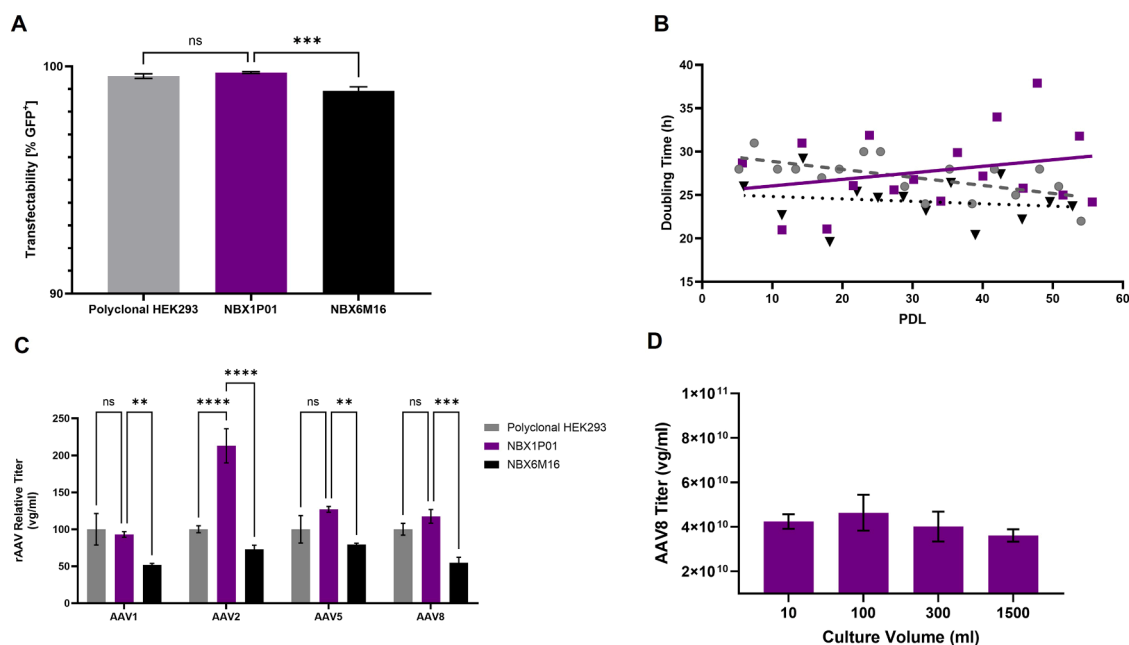


Fig. 2. Phenotypic characteristics of high-performing and low performing HEK293 cell clones. **A)** Transfectability of parental HEK293 cells and of the NBX1P01 and NBX6M16 cell subclones was measured by scoring the percentage of GFP⁺ cells following transient transfections of an rAAV-GFP plasmid ($N = 8$). **B)** Doubling time of the cells over 55 population doubling levels (PDL) of cultures of the parental cells (circles) and of the NBX1P01 (squares) or NBX6M16 (triangles) subclones. **C)** Vector titers were determined on crude rAAV preparations produced by NBX1P01 and NBX6M16 cells, relative to those produced by the parental HEK293 cells set to a value of 100 ($N = 3$). **D)** Mean rAAV8-EGFP crude titers obtained from increasing volumes of NBX1P01 cell suspension cultures ($N = 3$). Bars represent means; error bars indicate S.D. ns=not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

recombinant genome sizes, 8 ml production cultures were carried out to generate rAAV8 carrying combinations of promoters and transgenes ranging from 1.4 Kbp to 5.4 Kbp (Fig. 3A). Vector production titers decrease as the size of the cargo increased and there was a 10-fold difference between rAAV8-afibercept (2.4 Kbp) and rAAV8-BDD hF.VIII (5.4 Kbp). To demonstrate NBX1P01 capacity to produce rAAV of various serotypes with a range of transgenes, 8 mL production cultures were carried out to generate rAAV1-Nano-Luciferase, rAAV2-mRPE65 (the equivalent of Luxturna®), AAV5-BDD-hF.VIII (the equivalent of Roctavian®), rAAV8-EGFP, rAAV8-afibercept (the equivalent of ADVM-022) and AAV9-LacZ (Fig. 3B). Most crude vector titers obtained were around 10^{10} vg/mL with the notable exceptions of rAAV2-mRPE65 and AAV9-LacZ. AAV9 serotype 2 is known to be cell membrane associated and crude vector titers are expected to be lower than those of other serotypes. The LacZ cassette is a large transgene of 4.5 Kbp and crude titers were therefore also expected to be lower. These studies

demonstrated the robustness and versatility of NBX1P01 for the production of a variety of rAAV vectors comprising transgenes and serotypes of clinical relevance.

3.2. Comparison of NBX1P01 with a leading commercial HEK293 cell line

A series of head-to-head comparisons were carried out to evaluate the productivity of NBX1P01 against a leading commercial HEK293 rAAV production cell line using the same protocol for culture, transfection, harvest and quantification. Recombinant AAV8-EGFP and rAAV8-afibercept were produced in 10 mL cultures by the triple transfection method. Crude rAAV8-EGFP titers obtained from both cell lines were statistically similar ($p > 0.01$). However, rAAV8-afibercept titers produced by NBX1P01 were significantly higher than those from the commercial cell line ($p = 0.0266$, Fig. 4). In an additional study, 1.5

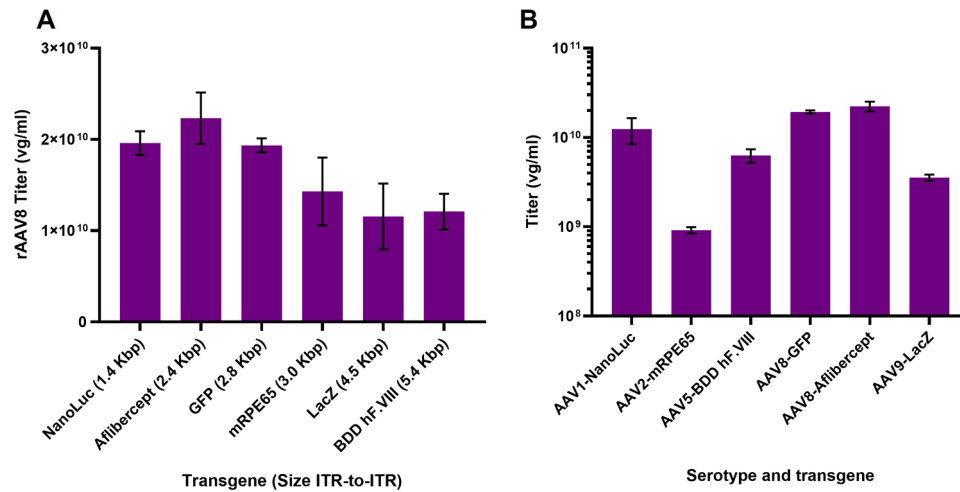


Fig. 3. Recombinant AAV production by NBX1P01 cells. **A)** rAA8 vectors were produced using NBX1P01 using various rAAV genome plasmids comprising representative promoter-transgene combinations of various lengths (Table 1). **B)** rAAV production of various serotypes and promoter-transgene combinations. Bars indicate means of at least 2 independent vector productions carried out in duplicate. Error bars indicate standard errors of the means.

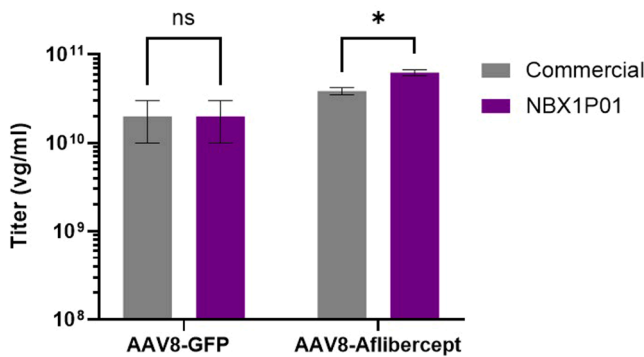


Fig. 4. Comparison of rAAV8 productions between the commercially available HEK293 cell line VPC2.0 and NBX1P01 using two different transgene cassettes: CMV-EGFP and CMV-aflibercept. Bars indicate mean crude titers from 3 vector productions at 10 mL scale, analyzed in duplicate; error bars indicate standard error of the means. * indicates $p = 0.0191$.

L vector productions of both rAAV8-EGFP and rAAV8-aflibercept were carried out and, following an affinity chromatography concentration step, rAAV8 full/empty ratios were evaluated by mass photometry. NBX1P01 produced 16 % of full rAAV8-EGFP particles whereas the commercial cell line produced 8 % full particles (Table 2). Similarly, NBX1P01 produced 23 % of full AAV8-aflibercept particles versus 8 % of full particles produced by the commercial cell line (Table 2). Lastly, long-read DNA sequencing using Oxford Nanopore Technologies (ONT) revealed that the quality and integrity of the vectors produced by both cell lines was comparable, both producing >70 % of intact genomes. The encapsidation frequency of Rep/Cap, Adv Helper, plasmid and HEK293 host DNA sequences was similar in vectors produced by either cell line.

Thus, vector production using NBX1P01 resulted in equal or more encapsidated EGFP or aflibercept sequences overall, when compared to vectors produced with the commercial cell line (Table 2). These studies demonstrated that the HP clone NBX1P01 is able to generate similar or higher amounts of rAAV than a leading commercial HEK293 cell line using the same protocol for culture, transfection, harvest and quantification, while consistently yielding higher full particle ratios and higher encapsidation frequencies of an intact gene of interest.

3.3. Genomic characterization of the NBX1P01 cell clone

Optical mapping and long-read sequencing were carried out to characterize both the HP and LP clone genomes and to compare them to the parental cell line. As expected, no evidence of Y-chromosome-derived sequences was observed using optical mapping (Fig. 5A) and the pseudotriploidy of the cells was also confirmed. Chromosomes 1, 2, 5, 7, 11, 12, 17, 21, 22 and X were observed to be triploid, while chromosomes 2, 3, 9, 14, 16, 18 and 20 were diploid in all three lines. However, some ploidy alterations were specifically observed in the LP clone. For instance, chromosomes 6 and 13 were observed to be tetraploid for both the parental cell line and the HP clone NBX1P01 but triploid in the LP clone NBX6M16. Interestingly, chromosome 10 was found to be diploid in the HP clone whereas it was triploid in both the parental and LP clone. Taking the human genome as reference, a genomic translocation analysis using optical mapping revealed 37 genomic translocation events for NBX1P01, versus the 50 and 61 translocation events of the NBX6M16 and parental HEK293 cells, respectively (Fig. 5B-D). The assembled genomes of these cells were searched for the E1A and E1B coding sequences, revealing the presence of Adv E1A/E1B at the expected PSG4 gene region of chromosome 19 (data not shown).

As previously reported, an approach for clonality assessment employs the usage of WGS data to calculate Variant Allele Frequencies (VAF) at putative haploid regions [17]; [28]). Therefore, a VAF analysis

Table 2

Summary of analytics obtained from rAAV8-EGFP and rAAV8-aflibercept produced in 1.5 L cultures using either a commercial HEK293 cell line or NBX1P01 and purified using affinity chromatography.

Vector*	Cell Line	Mass Photometry	Long Read DNA Sequencing				
		Full:Empty	Intact:Partial	% GOI	% Rep/Cap	% Helper	% Plasmid
AAV8-EGFP	Commercial	8:92	75:25	90.97	0.14	0.13	0.47
	NBX1P01	16:84	73:27	91.09	0.14	0.11	0.27
AAV8-Aflibercept	Commercial	8:92	77:23	90.95	0.14	0.14	0.75
	NBX1P01	23:72	77:23	91.43	0.11	0.11	0.3

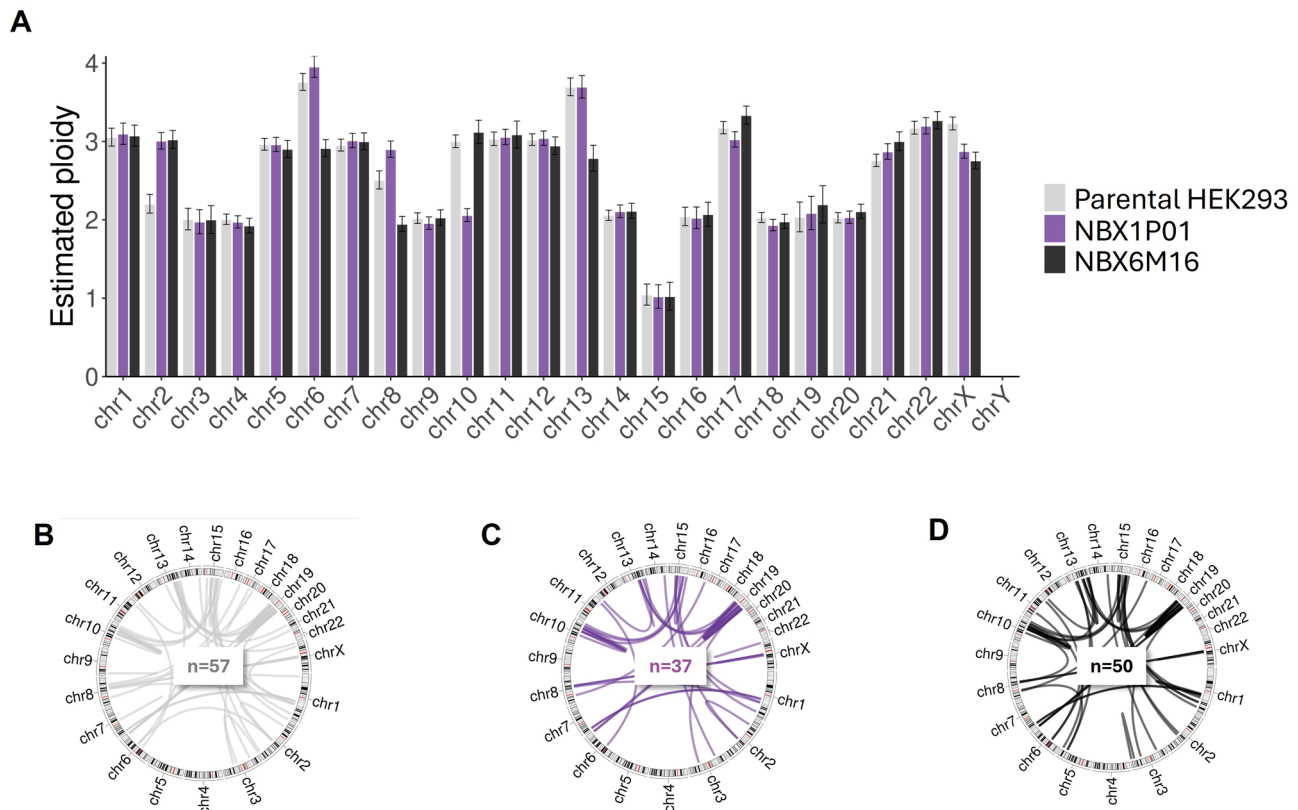


Fig. 5. Optical genome mapping (OGM) summary of the parental HEK293-FB cell line (gray), NBX1P01 (purple) and NBX6M16 (black). **A)** Bar graph showing the average number of each chromosome obtained from the three cell lines; error bars indicate twice the standard deviation of the molecule coverage for each chromosome. **B-D)** Genome-wide circos plots for each cell line showing the number of translocations detected for the three cell lines in relation to the human genome assembly (hg38).

was carried out on a haploid portion of chromosome 15 of the parental HEK293 line and of NBX1P01 and NBX6M16 cells. A uniform allele distribution along this region of the parental HEK293 cell line confirmed that this cell line is highly polyclonal (Fig. 6A). In contrast, enrichment of single nucleotide polymorphisms (SNPs) having VAF of 1 were observed in both NBX1P01 and NBX6M16 (Figs. 6C), indicating the clonality of these two cell lines. In addition, the number of single nucleotide variants (SNV, Fig. 6D) and structural variations (SV, i.e. insertions, deletions and translocations, Fig. 6E) using long-read sequencing and optical mapping were determined. Significantly fewer events were observed in both the HP and LP clones compared to the parental cell line, demonstrating a lower genomic heterogeneity. Interestingly, the HP clone was found to have a fewer number of both SNV (Fig. 6D) and SV (Fig. 6E) than the LP clone. Lastly, copy number variation (CNV) analysis was carried out to determine similarities and differences between the three cell lines, showing that the LP clone was more distantly related to either the parental cells or the HP clone (Fig. 7F). Both the parental and HP displayed a CN gain in the q arm of chromosome 13; the parental cell had a gain of the initial part of the p arm of chromosome 6, whereas the HP clone showed an extension of this CN gain to the rest of the chromosome (Fig. 7F). These results indicated that the HP NBX1P01 cells are genomically clonal and show lower genetic heterogeneity than the parental cell line and the LP clone.

3.4. Long-term genomic stability of NBX1P01

To determine the genomic stability over cell divisions of the NBX1P01 HP clone, the cells were passaged for 55 population doubling level (PDL) and analyzed again using optical mapping and long-range sequencing. Long-read sequencing revealed that the genome of NBX1P01 was 99 % identical between PDL 0 and PDL 55 cell

populations (data not shown), whereas few syntenic mutations, inversions, translocations and duplications were observed (Fig. 7A). Optical mapping analysis demonstrated that the genome-wide ploidy of NBX1P01 remained unmodified (Fig. 7B) after 55 PDL. AAV8-EGFP production was carried out to compare PDL0 and PDL55 cells, and no significant differences in vector production titers were observed ($p = 0.789$, Fig. 7C). These results indicated that the NBX1P01 HP clone is genetically and phenotypically stable over at least 55 PDL.

4. Discussion

Since their original description [1], HEK293 cells have been extensively used in bioproduction, specifically for viral vectors including lentiviral and recombinant AAV-based vectors. Different HEK293 lineages have been generated, propagated under different conditions, adapted to various growth media, selected to grow in either suspension or adherent platforms, cloned and subcloned to obtain desired phenotypes [9,29,30]. More recently, genetically modified HEK293T cells were reported for the production of rAAV [31]. Several biotherapeutics produced in HEK293 cells have been approved for human use by regulatory agencies [32]. Here, we describe the generation, selection and characterization of a HEK293 clonal cell line derived from a polyclonal source, following single cell cloning and phenotypic screening. The traditional method to generate single cell clones is through limiting dilution, however we chose to use single-cell printing with optical tracking to maintain traceability of the single cell clones. The first selection criteria were the single cell cloning survival and transfectability potential. HEK293 cells are relatively easy to transfect [33] and the cationic polymer polyethylenimine (PEI) is one of the most commonly used reagents for plasmid DNA transfection into HEK293 cells [34]. Our data showed that there were significant differences in transfection

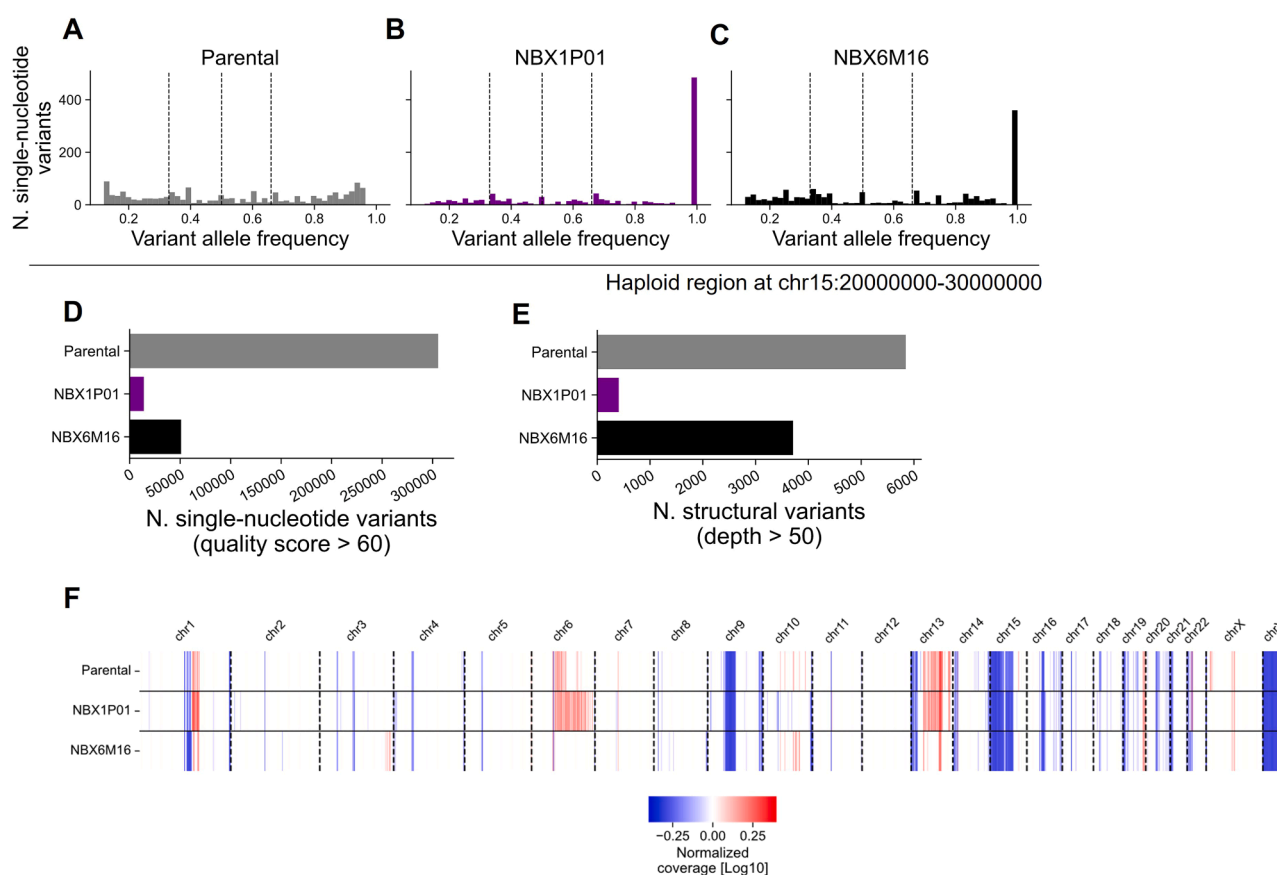


Fig. 6. Summary of the PacBio WGS-based genomic characterization of the parental HEK293-FB cell line (gray), NBX1P01 (purple) and NBX6M16 (brown). **A-C)** Determination of clonality based on variant allele frequency (VAF) in a region of chromosome 15 found to be haploid in the three cell lines. **D)** Number of bi-allelic single nucleotide variants (SNV) and **E)** structural variants (SV) detected in the three cell lines and not related to reference mismatch with the HEK293 genome. **F)** Copy number variations (CNV, increasing (red) or decreasing (blue)) across the genome of all three cell lines, visualized as Log₁₀-normalized read coverage in 1Mb genomic bins. SNP, SV and CNV calling were performed using the Telomere-to-Telomere human genome as a reference.

efficiencies between the tested clones (Supplementary Figure 1). One particular HP clone, NBX1P01, showed consistently high levels of transfection efficiency, whereas an LP clone, NBX6M16, showed consistently low levels of transfection efficiency. This suggested that there are genotypic or phenotypic differences in polyclonal populations of the same type that result in a differential capacity to uptake DNA and express the encoded recombinant proteins. The transfection efficiency of cells can be optimized depending on the transgene of interest ([9,35]; de Los Milagros Bassani Molinas et al., [36,32]). Thus, we used in this study a standardized, predictable and well-established transfection protocol as a first-pass approach [9,10].

A helper virus-free packaging system – also known as the triple plasmid transient transfection system – for the production of rAAV in HEK293 cells [10] is now the most commonly-used rAAV production system in research, clinical and commercial use [32]. This system was used to evaluate the capacity of the chosen HEK293 clones to produce rAAV of various serotypes. The HP clone outperformed the LP clone in AAV production when normalized to the parental HEK293 cell line. We did not investigate if the relatively low rAAV production by the LP clone 6M16 was related to its low transfectability or whether other biological factors could have influenced these observations. However, we found that the transfection efficacy of NBX1P01 clone was indistinguishable from that of the parental polyclonal population. Whether this stems from some of the observed genetic changes and/or from epigenetic alterations in NBX1P01 cells remains to be determined. Process development (PD) and design of experiment (DOE) are normally used to optimize production parameters and we have exploited both of these approaches to improve rAAV production in NBX1P01 achieving $>1 \times 10^{12}$ vg/ml of

rAAV8-GFP in crude supernatants (Kolcak-Yasli et al., Manuscript in preparation).

The capacity of NBX1P01 to produce consistent levels of rAAV in culture volumes ranging from 10 mL to 1.5 L was also demonstrated, as well as the capacity of NBX1P01 to produce rAAV carrying transgenes of different sizes, ranging from 1.4 Kbp to 5.4 Kbp. Vector productions with smaller transgenes resulted in higher production levels when compared to vectors carrying larger transgenes as observed previously [37,38], illustrating the ability of NBX1P01 cells to produce rAAV vectors of various transgene sizes. In a head-to-head comparison using the same protocol for culture, transfection, harvest and quantification, rAAV8 production with two different transgenes was carried out using NBX1P01 and a leading commercially-available HEK293 cell line, and we did not observe any statistically-significant differences between the vector titers.

Although vector titers are generally considered the main critical quality attribute (CQA) of rAAV production, both empty:full capsid ratios and packaged DNA analysis, including genome integrity and the presence of contaminating sequences, are becoming standard assessments of quality. For instance, mass photometry (MP) provides a fast, relatively low cost and accurate way of quantifying empty:full capsid ratios during vector production [39,11,37,12]. We observed that the production of full capsids by NBX1P01 were similar or higher than those produced by the commercial cells. A possible explanation for these differences may stem from the clonal selection of NBX1P01, when compared to the commercial cell line which may yield more heterogeneous vectors.

Lastly, a genomic characterization of the encapsidated DNA was

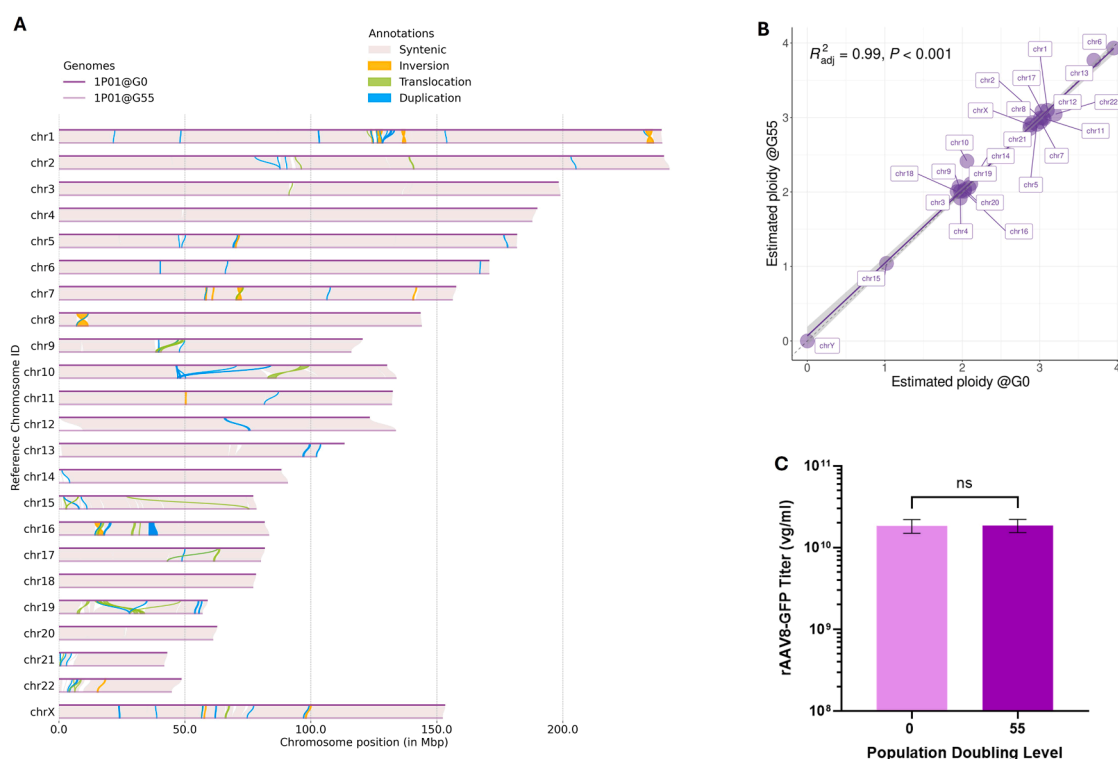


Fig. 7. Genomic and phenotypic stability of NBX1P01 between PDL0 and 55. **A)** Genomic stability determined by comparing the assemblies of NBX1P01 chromosomes over 55 generations. Specifically, we identified syntenic paths (grey), inversions (orange), translocations (green) and duplications (blue) events. **B)** Genomic stability of NBX1P01 determined by comparing the estimated ploidy at generation 0 (x axis) and 55 (y axis). **C)** Phenotypic stability determined by rAAV8 production by PDL0 and PDL55 cells. Bars indicate the means of three vector production processes quantified by duplicate measurements; error bars indicate standard errors of the means.

carried out using long-read and high-throughput DNA sequencing, commonly known as next generation sequencing (NGS). Sequencing approaches to characterize rAAV genome quality were developed, as based on a combination of Oxford Nanopore (ONT), Illumina, and PacificBio (PacBio) technologies, thereby providing much higher sensitivities and information than traditional PCR-based approaches [40–43]. Here, we found that the genomic quality, as defined by the percentage of intact genome and presence of contaminating sequences, of encapsidated vector DNA produced by NBX1P01 was comparable to that of the DNA produced by the leading commercial HEK293 cell line. The ratios of intact to partial genomes were similar in all samples analyzed, with >70 % intact rAAV genomes, similarly to some previous reports [44], but significantly higher than others [11]. The observed differences in the proportion of intact and partial AAV genomes may stem from the variations in purification methods used in this study versus others, or in the use of different quantification approaches, such as long-read sequencing technologies, digital or droplet PCR-based analytics. In this study, the vectors underwent only affinity chromatography enrichment, in contrast to other studies utilizing density ultracentrifugation for purification. Thus, the purification method, specifically the choice of downstream processing, may influence the quality and integrity of the rAAV packaged DNA. Head-to-head downstream process analysis aided by drug product potency assays will help to determine how different processes affect the quality of the therapeutic products.

HEK293 cells have been shown to have a high degree of genomic instability, displaying a pseudotriploid genotype and with frequent chromosomal rearrangements [7,8,45]. Similarly, CHO cells, which are also commonly used in biomanufacturing, have also been shown to be genetically unstable [46,47]. Using optical genome mapping (OGM), we demonstrated that NBX1P01 retains the pseudotriploidy characteristic of HEK293 cells. Comparing the HEK293 with the human genome

assembly, we found the parental HEK293 polyclonal cell line contained 61 structural variations, whereas the LP clone contained 50 and the HP clone NBX1P01 only 37. This indicates a lower degree of genetic heterogeneity of NBX1P01 when compared to the parental cell line and most importantly, this level of stability was demonstrated over 55 PDL, which correlated with the ability of NBX1P01 to consistently produce high-levels of rAAV after continuous passaging.

The inserted Adv5 E1A/B sequences within HEK293 cells were originally described to be within the pregnancy-specific β -1 glycoprotein 4 (PSG4) gene as a single, 4344 nucleotide collinear insertion of viral DNA without any rearrangements [48]. It has been previously reported that the copies of E1A/B can vary between 5 and 6 depending on the HEK293 line [7] with some reports indicating the presence of up to 15 copies [49]. Because Adv E1A/B have been proposed to be necessary for the optimal production of rAAV [50,51], the copy number and resulting differential expression of these viral proteins may have a direct impact on rAAV production. During our studies, we calculated that the original HEK293 cell line from which our clones were derived had three copies of E1A/B and that the HP clone NBX1P01 also contained the three copies in the same locus. This data further supports the genomic stability profile of NBX1P01.

As previously reported, an approach for clonality assessment employs the usage of WGS data to calculate Variant Allele Frequencies (VAF) at putative haploid regions [28,17]. Our studies demonstrated that both NBX1P01 and NBX6M16 are clonal populations, as observed by the high number of VAF equal to 1 within the haploid region of chromosome 15, in contrast to the VAF observed for the parental HEK293 cell line. Single nucleotide variant (SNV) and structural variant (SV) quantification revealed that NBX1P01 contains significantly fewer variations than both the parental HEK293 or the LP clone NBX6M16. SNV and SV are commonly used in cancer genomic studies to determine clonality origin of oncogenesis and other cytogenetic profiles [52–54]

and SNV-based analytical processes were developed to demonstrate the monoclonality of CHO cell lines [17]. As biomanufacturing progresses, the unstable nature of the cell lines used in manufacturing becomes even more relevant, and choosing the most genomically-stable clones has been shown to improve manufacturing yields [55].

It has become clear that many HEK293 cell lines available today are very heterogeneous, containing cells displaying highly variable genotypes and phenotypes [49,7,8]. This heterogeneity represents a significant disadvantage during biomanufacturing, resulting in differences in transfection efficiencies and batch-to-batch variability in cell culture performance. It is now well-recognized that there should be an emphasis on generating genetically homogeneous and stable cell lines, in addition to high-production abilities. The favored approach is to carry out single-cell cloning followed by extensive genomic and functional characterization of the resulting clones. Although this approach originally creates well-defined subcultures with optimal performance, the unstable nature of the cells tends to result in genetic and functional variations over time [56,57]. Overall, using stable as well as high-yield HEK293 cell lines is the ultimate goal in rAAV manufacturing, as it would reduce costs of manufacturing and the occurrence of failed batches, while increasing the quality of the drug substance [58,13]. Ongoing transient transfection optimization studies in NBX1P01 cells have yielded crude rAAV titers above 10^{12} vg/ml, indicating that very high productivities may be achieved with these cells (unpublished results). Therefore, using stable and genomically homogeneous cell clones to produce biologics by transient processes may be a significant step towards achieving this goal.

Declaration of Competing Interest

The authors are currently employed and/or own share or options of NewBiologix SA.

CRediT authorship contribution statement

Mélissa Vona: Investigation. **Iris Bodenmann:** Investigation. **Marc-Antoine Perrenoud:** Investigation. **Rachel Buchs:** Investigation. **Pelin Kolcak Yasli:** Investigation. **Luca Nanni:** Investigation. **Romain Daveau:** Investigation. **Alexandre Félix:** Investigation. **Jens Stolte:** Investigation. **Ann-Kristin Hov:** Investigation. **Bertrand Chollet:** Investigation. **Thierry Schuepbach:** Investigation. **Déborah Ley:** Investigation. **Efrain Guzman:** Writing – review & editing, Writing – original draft, Supervision, Project administration. **Igor Fisch:** Writing – review & editing, Funding acquisition, Conceptualization. **Nicolas Mermod:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Efrain Guzman reports a relationship with NewBiologix SA that includes: employment. Mélissa Vona reports a relationship with NewBiologix SA that includes: employment. Iris Bodenmann reports a relationship with NewBiologix SA that includes: employment. Marc-Antoine Perrenoud reports a relationship with NewBiologix SA that includes: employment. Rachel Buchs reports a relationship with NewBiologix SA that includes: employment. Pelin Kolcak Yasli reports a relationship with NewBiologix SA that includes: employment. Luca Nanni reports a relationship with NewBiologix SA that includes: employment. Romain Daveau reports a relationship with NewBiologix SA that includes: employment. Alexandre Felix reports a relationship with NewBiologix SA that includes: employment. Jens Stolte reports a relationship with NewBiologix SA that includes: employment. Ann-Kristin Hov reports a relationship with NewBiologix SA that includes: employment. Bertrand Chollet reports a relationship with NewBiologix SA that includes: employment. Thierry Schuepbach reports a relationship with NewBiologix SA that includes: employment. Deborah Ley reports a relationship with NewBiologix SA that includes: employment.

Igor Fisch reports a relationship with NewBiologix SA that includes: board membership, employment, and equity or stocks. Nicolas Mermod reports a relationship with NewBiologix SA that includes: board membership, employment, and equity or stocks. N/A If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank all members of NewBiologix for their continued support and scientific discussion, Valerie Le Fourn and Pierre-Olivier Duroy for their contributions to this work. This work was fully financed by NewBiologix SA.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.btre.2026.e00948](https://doi.org/10.1016/j.btre.2026.e00948).

Data availability

Data will be made available on request.

References

- [1] W.C. Russell, et al., Characteristics of a Human cell line transformed by DNA from human adenovirus type 5', *J. Gen. Virol.* 36 (1) (1977) 59–72, <https://doi.org/10.1099/0022-1317-36-1-59>. Available at.
- [2] J. Yuan, et al., The scattered twelve tribes of HEK293', *Biomed. Pharmacol. J.* 11 (2) (2018) 621–623, <https://doi.org/10.13005/bpj/1414>. Available at.
- [3] E. Casademunt, et al., The first recombinant human coagulation factor VIII of human origin: human cell line and manufacturing characteristics, *Eur. J. Haematol.* 89 (2) (2012) 165–176, <https://doi.org/10.1111/j.1600-0609.2012.01804.x>. Available at.
- [4] C. Kannicht, et al., Characterisation of the post-translational modifications of a novel, human cell line-derived recombinant human factor VIII', *Thromb. Res.* 131 (1) (2013) 78–88, <https://doi.org/10.1016/j.thromres.2012.09.011>. Available at.
- [5] W.D. Richardson, H. Westphal, 'Requirement for either early region 1a or early region 1b adenovirus gene products in the helper effect for adeno-associated virus', *J. Virol.* 51 (2) (1984) 404–410, <https://doi.org/10.1128/jvi.51.2.404-410.1984>. Available at.
- [6] M.F. Naso, et al., Adeno-associated virus (AAV) as a vector for gene therapy, *BioDrugs* 31 (4) (2017) 317–334, <https://doi.org/10.1007/s40259-017-0234-5>. Available at.
- [7] Y.C. Lin, et al., Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations, *Nat. Commun.* 5 (1) (2014) 4767, <https://doi.org/10.1038/ncomms5767>. Available at.
- [8] M. Malm, et al., Evolution from adherent to suspension: systems biology of HEK293 cell line development, *Sci. Rep.* 10 (1) (2020) 18996, <https://doi.org/10.1038/s41598-020-76137-8>. Available at.
- [9] P.S. Chahal, et al., Production of adeno-associated virus (AAV) serotypes by transient transfection of HEK293 cell suspension cultures for gene delivery, *J. Virol. Methods* 196 (2014) 163–173, <https://doi.org/10.1016/j.jviromet.2013.10.038>. Available at.
- [10] R.F. Collaco, X. Cao, J.P. Trempe, A helper virus-free packaging system for recombinant adeno-associated virus vectors, *Gene* 238 (2) (1999) 397–405, [https://doi.org/10.1016/s0378-1119\(99\)00347-9](https://doi.org/10.1016/s0378-1119(99)00347-9). Available at.
- [11] P. Eisenhut, et al., Orthogonal characterization of rAAV9 reveals unexpected transgene heterogeneity', *J. Biotechnol.* 393 (2024) 128–139, <https://doi.org/10.1016/j.jbiotec.2024.07.020>. Available at.
- [12] C. Wagner, et al., Quantification of empty, partially filled and full adeno-associated virus vectors using mass photometry, *Int. J. Mol. Sci.* 24 (13) (2023) 11033, <https://doi.org/10.3390/ijms241311033>. Available at.
- [13] J. Zhang, et al., Thorough molecular configuration analysis of noncanonical AAV genomes in AAV vector preparations, *Mol. Ther. - Methods Clin. Dev.* 32 (1) (2024) 101215, <https://doi.org/10.1016/j.omtm.2024.101215>. Available at.
- [14] S. Nurk, et al., The complete sequence of a human genome, *Science* (1979) 376 (6588) (2022) 44–53, <https://doi.org/10.1126/science.abj6987>. Available at.
- [15] H. Li, 'Minimap2: pairwise alignment for nucleotide sequences', *Bioinform. Ed. I. Birol.* 34 (18) (2018) 3094–3100, <https://doi.org/10.1093/bioinformatics/bty191>. Available at.
- [16] R. Poplin, et al., A universal SNP and small-indel variant caller using deep neural networks, *Nat. Biotechnol.* 36 (10) (2018) 983–987, <https://doi.org/10.1038/nbt.4235>. Available at.
- [17] A. Kuhn, et al., Genome-wide analysis of single nucleotide variants allows for robust and accurate assessment of clonal derivation in cell lines used to produce

- biologics, *Biotechnol. Bioeng.* 117 (12) (2020) 3628–3638, <https://doi.org/10.1002/bit.27534>. Available at.
- [18] H. Cheng, et al., Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm, *Nat. Methods* 18 (2) (2021) 170–175, <https://doi.org/10.1038/s41592-020-01056-5>. Available at.
- [19] M. Alonge, et al., Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing, *Genome Biol.* 23 (1) (2022) 258, <https://doi.org/10.1186/s13059-022-02823-7>. Available at.
- [20] F. Ramírez, et al., 'deepTools: a flexible platform for exploring deep-sequencing data', *Nucleic. Acids. Res.* 42 (W1) (2014) W187–W191, <https://doi.org/10.1093/nar/gku365>. Available at.
- [21] C.L. Weiß, et al., nQuire: a statistical framework for ploidy estimation using next generation sequencing', *BMC Bioinform.* 19 (1) (2018) 122, <https://doi.org/10.1186/s12859-018-2128-z>. Available at.
- [22] A. Morgulis, et al., WindowMasker: window-based masker for sequenced genomes', *Bioinformatics* 22 (2) (2006) 134–141, <https://doi.org/10.1093/bioinformatics/bti774>. Available at.
- [23] M. Goel, et al., SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies, *Genome Biol.* 20 (1) (2019) 277, <https://doi.org/10.1186/s13059-019-1911-0>. Available at.
- [24] A. Shumate, S.L. Salzberg, Liftoff: accurate mapping of gene annotations, *Bioinformatics* 37 (12) (2021) 1639–1643, <https://doi.org/10.1093/bioinformatics/btaa1016>. Edited by A. Valencia Available at.
- [25] S. Chen, et al., fastp: an ultra-fast all-in-one FASTQ preprocessor, *Bioinformatics* 34 (17) (2018) i884–i890, <https://doi.org/10.1093/bioinformatics/bty560>. Available at.
- [26] A. Rhie, et al., Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies, *Genome Biol.* 21 (1) (2020) 245, <https://doi.org/10.1186/s13059-020-02134-9>. Available at.
- [27] M. Seppey, M. Manni, E.M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness, in: M. Kollmar (Ed.), *Gene Prediction*, NY: Springer New York (Methods in Molecular Biology), New York, 2019, pp. 227–245, https://doi.org/10.1007/978-1-4939-9173-0_14. Available at.
- [28] L. Boscolo Bielo, et al., Variant allele frequency: a decision-making tool in precision oncology? *Trends. Cancer* 9 (12) (2023) 1058–1068, <https://doi.org/10.1016/j.trecan.2023.08.011>. Available at.
- [29] M. Lu, et al., Enhancing the production of recombinant adeno-associated virus in synthetic cell lines through systematic characterization, *Biotechnol. Bioeng.* 121 (1) (2024) 341–354, <https://doi.org/10.1002/bit.28562>. Available at.
- [30] S. Satkunanathan, et al., 'Establishment of a novel cell line for the enhanced production of recombinant adeno-associated virus vectors for gene therapy', *Hum. Gene Ther.* 25 (11) (2014) 929–941, <https://doi.org/10.1089/hum.2014.041>. Available at.
- [31] C. Croissant, et al., A new T-antigen negative HEK293 cell line with improved AAV productivity', *Biotechnol. Bioeng.* 120 (7) (2023) 1953–1960, <https://doi.org/10.1002/bit.28414>. Available at.
- [32] E. Tan, et al., HEK293 Cell line as a platform to produce recombinant proteins and viral vectors, *Front. Bioeng. Biotechnol.* 9 (2021) 796991, <https://doi.org/10.3389/fbioe.2021.796991>. Available at.
- [33] L. Porosk, et al., Predicting transiently expressed protein yields: comparison of transfection methods in CHO and HEK293, *Pharmaceutics* 14 (9) (2022) 1949, <https://doi.org/10.3390/pharmaceutics14091949>. Available at.
- [34] O. Boussif, et al., A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine, *Proc. Natl. Acad. Sci. U.S.A.* 92 (16) (1995) 7297–7301, <https://doi.org/10.1073/pnas.92.16.7297>. Available at.
- [35] K. Swiech, et al., Transient transfection of serum-free suspension HEK 293 cell culture for efficient production of human rFVIII', *BMC. Biotechnol.* 11 (1) (2011) 114, <https://doi.org/10.1186/1472-6750-11-114>. Available at.
- [36] Milagros de Los, M. Bassani Molinas, et al., Optimizing the transient transfection process of HEK-293 suspension cells for protein production by nucleotide ratio monitoring, *Cytotechnology* 66 (3) (2014) 493–514, <https://doi.org/10.1007/s10616-013-9601-3>. Available at.
- [37] C. Hiemenz, et al., Genome length determination in adeno-associated virus vectors with mass photometry', *Mol. Ther. - Methods Clin. Dev.* 31 (2023) 101162 <https://doi.org/10.1016/j.omtm.2023.101162>. Available at.
- [38] N. Ibreljic, B.E. Draper, C.W. Lawton, Recombinant AAV genome size effect on viral vector production, purification, and thermostability, *Mol. Ther. - Methods Clin. Dev.* 32 (1) (2024) 101188, <https://doi.org/10.1016/j.omtm.2024.101188>. Available at.
- [39] L.F. Barnes, et al., Quantitative analysis of genome packaging in recombinant AAV vectors by charge detection mass spectrometry', *Mol. Ther. - Methods Clin. Dev.* 23 (2021) 87–97, <https://doi.org/10.1016/j.omtm.2021.08.002>. Available at.
- [40] M.A. Brimble, S.M. Winston, A.M. Davidoff, Stowaways in the cargo: contaminating nucleic acids in rAAV preparations for gene therapy, *Mol. Ther.* 31 (10) (2023) 2826–2838, <https://doi.org/10.1016/j.ymthe.2023.07.025>. Available at.
- [41] K. Guerin, et al., 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 (11–12) (2020) 664–678, <https://doi.org/10.1089/hum.2019.277>. Available at.
- [42] S. Namkung, et al., Direct ITR-to-ITR nanopore sequencing of AAV vector genomes, *Hum. Gene Ther.* (2022), <https://doi.org/10.1089/hum.2022.143> p. hum.2022.143. Available at.
- [43] P.W.L. Tai, et al., Adeno-associated virus genome population sequencing achieves full vector genome resolution and reveals Human-vector chimeras, *Mol. Ther. - Methods Clin. Dev.* 9 (2018) 130–141, <https://doi.org/10.1016/j.omtm.2018.02.002>. Available at.
- [44] T. Duong, et al., Three-dimensional linkage analysis with digital PCR for genome integrity and identity of recombinant adeno-associated virus, *Sci. Rep.* 15 (1) (2025) 2154, <https://doi.org/10.1038/s41598-024-77378-7>. Available at.
- [45] A.A. Stepanenko, V.V. Dmitrenko, HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution, *Gene* 569 (2) (2015) 182–190, <https://doi.org/10.1016/j.gene.2015.05.065>. Available at.
- [46] S. Hammond, et al., Genomic sequencing and analysis of a Chinese hamster ovary cell line using Illumina sequencing technology, *BMC Genom.* 12 (1) (2011) 67, <https://doi.org/10.1186/1471-2164-12-67>. Available at.
- [47] Wurm, F. (2013) 'CHO quaspecies—Implications for manufacturing processes', *Processes*, 1(3), pp. 296–311. Available at: <https://doi.org/10.3390/pr1030296>.
- [48] N. Louis, C. Eveleigh, F.L. Graham, Cloning and sequencing of the cellular–Viral junctions from the Human Adenovirus type 5 transformed 293 cell line, *Virology* 233 (2) (1997) 423–429, <https://doi.org/10.1006/viro.1997.8597>. Available at.
- [49] X. Lian, et al., A new HEK293 cell with CR2 region of E1A gene deletion prevents the emergence of replication-competent Adenovirus, *Cancers* 15 (24) (2023) 5713, <https://doi.org/10.3390/cancers15245713>. Available at.
- [50] D. Catalán-Tatjer, et al., Unravelling the essential elements for recombinant adeno-associated virus (rAAV) production in animal cell-based platforms, *Biotechnol. Adv.* 73 (2024) 108370, <https://doi.org/10.1016/j.biotechadv.2024.108370>. Available at.
- [51] T. Matsushita, et al., 'The adenovirus E1A and E1B19K genes provide a helper function for transfection-based adeno-associated virus vector production', *J. Gen. Virol.* 85 (8) (2004) 2209–2214, <https://doi.org/10.1099/vir.0.79940-0>. Available at.
- [52] A. Chakraborty, F. Ay, Identification of copy number variations and translocations in cancer cells from Hi-C data, *Bioinformatics* 34 (2) (2018) 338–345, <https://doi.org/10.1093/bioinformatics/btx664>. Edited by C. Curtis Available at.
- [53] F. Sallustio, et al., Identification and monitoring of copy number variants (CNV) in monoclonal gammopathy, *Cancer Biol. Ther.* 22 (5–6) (2021) 404–412, <https://doi.org/10.1080/15384047.2021.1946458>. Available at.
- [54] X. Shao, et al., Copy number variation is highly correlated with differential gene expression: a pan-cancer study, *BMC. Med. Genet.* 20 (1) (2019) 175, <https://doi.org/10.1186/s12881-019-0909-5>. Available at.
- [55] Q. He, et al., Improved titer in late-stage mammalian cell culture manufacturing by re-cloning, *Bioengineering* 9 (4) (2022) 173, <https://doi.org/10.3390/bioengineering9040173>. Available at.
- [56] S.A. O'Brien, et al., Multiplexed clonality verification of cell lines for protein biologic production, *Biotechnol. Prog.* 36 (4) (2020) e2978, <https://doi.org/10.1002/btpr.2978>. Available at.
- [57] S. Vcelar, et al., Changes in chromosome counts and patterns in CHO cell lines upon generation of recombinant cell lines and subcloning, *Biotechnol. J.* 13 (3) (2018) 1700495, <https://doi.org/10.1002/biot.201700495>. Available at.
- [58] S.G.R. Gurazada, et al., HEK-omics: the promise of omics to optimize HEK293 for recombinant adeno-associated virus (rAAV) gene therapy manufacturing, *Biotechnol. Adv.* 79 (2025) 108506, <https://doi.org/10.1016/j.biotechadv.2024.108506>. Available at.