Molecular Therapy
Methods & Clinical Development
Original Article

# Development and validation of a model gene therapy biodistribution assay for AVGN7 using digital droplet polymerase chain reaction

Buel D. Rodgers,[1] Sarah K. Herring,[2] Dereck R. Carias,[3] Joyce Chen,[3] and Agostinho G. Rocha[3]

[1]AAVogen, Rockville, MD 20850, USA; [2]Latham BioPharm Group, Elkridge, MD 21075, USA; [3]Smithers Pharmaceutical Development Services, Ewing, NJ 08628, USA

**Biodistribution assays are integral to gene therapy commercialization and have traditionally used real-time qPCR. Droplet digital PCR (ddPCR), however, has distinct advantages including higher sensitivity and absolute quantification but is underused because of lacking regulatory guidance and meaningful examples in the literature. We report a fit-for-purpose model process to validate a good laboratory practice (GLP)-compliant ddPCR assay for AVGN7, a Smad7 gene therapeutic for muscle wasting. Duplexed primer/probe sets for Smad7 and mouse TATA-box binding protein were optimized using gBlock DNA over a dynamic range of 10–80,000 copies/reaction in 250 ng mouse gDNA. Linearized plasmid and mouse gDNA were used for validation, which determined precision, accuracy, ruggedness/robustness, selectivity, recovery, specificity, dilution linearity, and stability. Inter-run precision and accuracy met previously established criteria with bias between −5% and 15%, coefficient of variation (CV) less than 19%, and total error within 8%–35%. The limit of detection was 2.5 copies/reaction, linearity was confirmed at 40–80,000 copies/reaction, specificity was demonstrated by single droplet populations and assay stability was demonstrated for benchtop, refrigerated storage, and repeated freeze-thaw cycles. The procedural road map provided exceeds recently established standards. It is also relevant to many IND-enabling processes, as validated ddPCR assays can be used in biodistribution studies and with vector titering and manufacturing quality control.**

## INTRODUCTION

Biodistribution assessments are required components of investigational and final drug or biological approval applications with the U.S. Food and Drug Administration (FDA), the European Medicines Agency (EMA), the Japan Pharmaceutical and Medical Devices Agency (PMDA), and other regulatory authorities.[1–4] During the nonclinical phase, they are often performed as components of small and large animal toxicology studies in compliance with good laboratory practice (GLP) and must be developed and, at a minimum, qualified or potentially validated in a manner consistent with the study. Validation is not necessarily required unless the biodistribution analysis is performed on samples collected from a study conducted under GLP compliance or if the assay will be used in clinical trials or in manufacturing clinical products, as for example with quality control (QC) or titering, which requires compliance with good manufacturing practice (GMP). This is generally the case for gene therapy programs that have traditionally relied upon real-time qPCR as the method of choice.

Droplet digital PCR (ddPCR) is an alternative to qPCR with many significant advantages.[5,6] These include un-paralleled analytical sensitivity, improved precision and accuracy, better reliability, resistance to inhibitors (e.g., detergents, heparin) and co-purified biological compounds (e.g., heme, urea) as well as absolute quantification without standard curves. Amplification occurs within aqueous-oil emulsion droplets where target template is partitioned according to a Poisson distribution. Endpoint, rather than threshold, fluorescent readouts are then collected to distinguish positive- from negative-reacting droplets. The fraction of positive droplets/partitions is then used in Poisson modeling to determine the concentration of target genes.[7] The heightened sensitivity of ddPCR is being exploited in diagnostic assays for different pathogens including SARS-CoV-2,[8,9] and at least one study has demonstrated superiority over qPCR in titering assays for self-complementary adeno-associated viral vectors.[10] Thus, the technology appears uniquely suited for gene therapy programs where highly sensitive pharmacokinetic, shedding, and biodistribution analyses take the place of conventional absorption, distribution, metabolism, and elimination (ADME) studies.[1] To date, however, the FDA has published only generalized guidance documents for bioanalytical method validation without specific mention of ddPCR,[2] and very few relevant examples can be found in the scientific literature. Recent reviews recognize this gap[5,11] and have supplemented the knowledge database with high level summaries of assay and regulatory requirements. Most notable is a white paper generated by the Global CRO Council in Bioanalysis (GCC) with authors from 26 different contract research organizations. This paper provides a consensus on validation parameters for relevant PCR-based assays, a "harmonized approach" to validation and a description of standard operating procedures (SOPs)[12] that follow Minimum Information for

Received 10 January 2023; accepted 8 May 2023;
https://doi.org/10.1016/j.omtm.2023.05.007.

**Correspondence:** Buel D. Rodgers, PhD, AAVogen, Inc., 13420 Glen Lea Way, Rockville, MD 20850, USA.
**E-mail:** danrodgers@aavogen.com

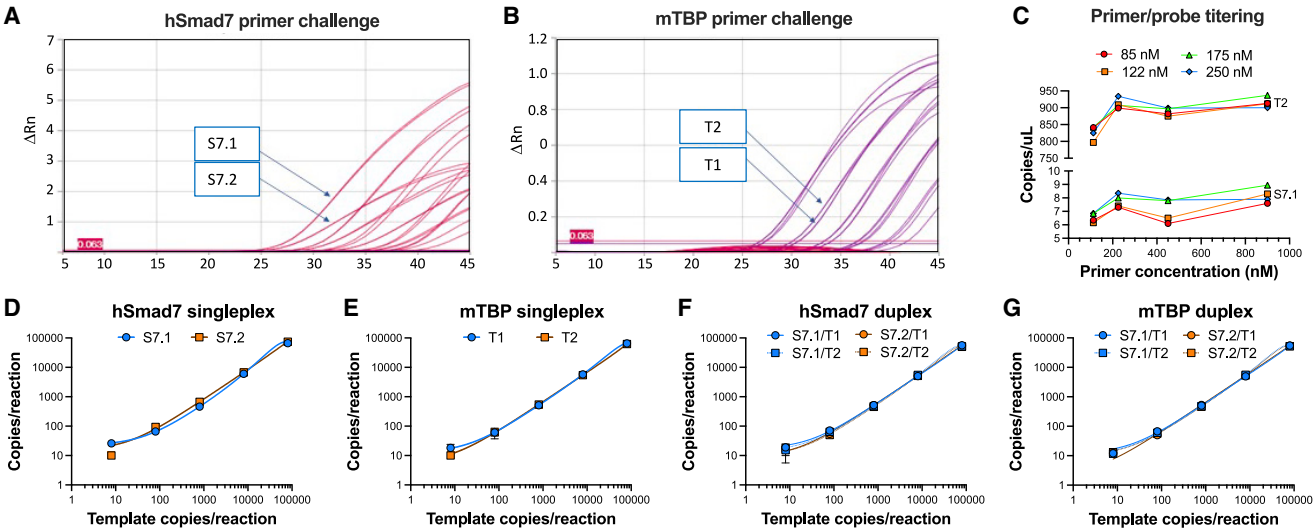


**Figure 1. Primer/probe characterization**

(A and B) Five 10-fold dilutions of synthetic gBlock DNA reference standards containing both hSmad7 and mTBP sequences were separately amplified using qPCR with the indicated primer (900 nM)/probe (250 nM) sets (Table 3). Duplicate samples were amplified for each set and horizontal red lines indicate the cycle threshold ($C_t$) for each curve. (C) Different concentrations of primers (x axis) and probes (key) were titered and used to amplify gBlock within expected assay concentrations of mouse gDNA and vector genomes. Titering assays were performed for all primer/probe sets (Table 4), but only T2 and S7.1 are shown. (D–G) Linearity assessment of the indicated primer/probe sets in singleplex and duplex reactions. gBlock template was diluted linearly as indicated and the data lines shown were fit using a third order polynomial.

Publication of Quantitative Real-Time PCR Experiments (MIQE)[13] and digital MIQE guidelines.[14]

Despite the dearth of available guidance, we recognize the power and overall benefit of the ddPCR approach and developed and validated an assay suitable for quantifying AVGN7, a gene therapeutic featuring the serotype 6 adeno-associated viral (AAV6) capsid and a human Smad7 (hSmad7) expression construct.[15] AVGN7 increases muscle mass and enhances muscle function in healthy mice and in different murine models of muscle wasting that include inflammatory insult and cancer cachexia.[16,17] It can be used with local or systemic administration and because it broadly attenuates multiple extracellular signals responsible for muscle wasting, has the potential to broadly treat a variety of muscle wasting conditions.[15] Expression of hSmad7 is tightly controlled by the CK8 muscle-specific promoter,[18–20] which should prevent significant off-target expression in non-muscle tissues. Several high-profile toxicities, however, have been reported for AAV8- and AAV9-based therapeutics, all of which appear to be associated with cytotoxic T cell responses in the liver.[21–23] Thus, biodistribution rather than transgene expression *per se* is a significant concern for most AAV-based gene therapy programs and underscores the importance of optimizing and validating highly sensitive assays.

Described herein is the development, optimization, and validation of a ddPCR assay for quantifying AVGN7 vector genomes in tissue extracts and whole blood. Our report follows the white paper published by the GCC[12] and outlines the critical processes within a typical commercial timeline. It describes assay performance, acceptance criteria, and metrics expected from several regulatory authorities. It also addresses challenges to the technology and serves as a model for guiding development of other gene therapy programs.

## RESULTS

### Assay development

Basic Local Alignment Search Tool (BLAST) analysis of all primers and probes confirmed 100% sequence homology to the targeted sequences without any evidence of significant homology for non-targeted sequences. Notwithstanding, preliminary studies using qPCR to amplify different concentrations of gBlock template, with the highest concentrations of primers (900 nM) and probes (250 nM) tested, identified significant differences in the hSmad7 primer/probe sets, but not in the mouse TATA-box binding protein (mTBP) sets. Such differences were readily apparent in the amplification curves as the signal plateau of S7.1 was almost double that of S7.2 (Figure 1A). The mTBP2 set produced a slightly earlier cycle threshold ($C_t$) than mTBP1 (Figure 1B), which could indicate better sensitivity, although this was not confirmed.

Titering experiments revealed superior performance for the highest concentrations of primers and probes for both hSmad7 and mTBP as variability in the observed template concentrations decreased as primer/probe concentrations increased (Figure 1C). Singleplex reactions using 900 nM primers and 250 nM probes were then used to quantify 10-fold dilutions of gBlock in comparison tests of assay accuracy, although no differences were noted (Figures 1D and 1E). Tests of primer compatibility in different permutations of duplex reactions also revealed no differences (Figures 1F and 1G).

**Table 1. Assay development: precision (% CV) and accuracy (% Bias)**

| Assay | Nominal copies per reaction | Observed copies per reaction | % CV | % Bias |
|---|---|---|---|---|
| HSMAD7 | 80,000 | 78,366 | 2.0 | −2.0 |
| | 40,000 | 38,906 | 2.6 | −2.7 |
| | 20,000 | 19,916 | 4.2 | −0.4 |
| | 2,000 | 2,000 | 4.0 | 0.0 |
| | 200 | 228 | 11.8 | 13.7 |
| | 20 | 30 | 16.5 | 46.7 |
| | 2 | 2 | 49.3 | 13.3 |
| MTBP | 80,000 | 79,334 | 1.8 | −0.8 |
| | 40,000 | 39,650 | 2.6 | −0.9 |
| | 20,000 | 20,304 | 4.0 | 1.5 |
| | 2,000 | 2,036 | 3.9 | 1.8 |
| | 200 | 230 | 11.2 | 14.7 |
| | 20 | 30 | 12.8 | 50.0 |
| | 2 | 2 | 49.3 | 13.3 |

Mean droplets/well, 17,899 ± 65 (SEM); copies/droplet, 0.0001–4.3.

Lower than expected plateaus in qPCR assays, as with S7.2, can potentially affect endpoint calculations. It is also frequently caused by limited primer and probe concentrations yet our preliminary assays used high concentrations of each. Although both primer/probe sets performed equally well in ddPCR singleplex and duplex assays, we eventually decided to focus on the S7.1 and T2 pairing to avoid the potential endpoint complications of S7.2 and because T1 primers mapped to different exons, producing a larger than expected amplicon. This latter effect could potentially influence assay accuracy when starting template concentrations are high and reagents become limiting in later amplification cycles.

Assay development of singleplex reactions containing 900 nM primers with 250 nM probes for S7.1 and T2 sets revealed similar levels of precision (% coefficient of variation [CV]) and accuracy (% bias) (Table 1). In fact, both measures for both primer/probe sets were less than or much less than 15% when quantifying 200–80,000 copies/reaction. Note that mean droplet numbers (17,163–18,361) and copies per droplet for each concentration tested (from 4.3 for 80,000 to 0.0001 for 2) remained well above and below, respectively, the manufacturer's recommended parameters of 10,000 and 5.0, which avoids overloading. These preliminary assessments also suggest that the limit of detection (LOD) falls to fewer than 10 copies/reaction, possibly between 2 and 20, as either the % CV or % bias, but not both, met our LOD validation criteria for both assays.

### Assay validation

Tests for accuracy (% bias) and precision (% CV) passed the acceptance criteria for all QCs and by large margins (Figure 2). This includes measures for both intra- and inter-assay variation, which separately assess robustness. It also includes assays for both genes, although the mTBP assay measured a fixed concentration (79,870 copies/reaction) for every hSmad7 QC level as this design replicates conditions of gene therapy biodistribution assays. There was a trend to underestimate hSmad7 levels on the high end of the QC scale and to overestimate on the low end. This effect was minimal, however, and below ±5% bias from low QC (LQC) to upper limit of quantitation (ULOQ). The greatest variability occurred with lower limit of quantitation (LLOQ) and was reflected in % CV and % bias both within and between assays (Figure 2E). Such findings were expected and typically occur with LLOQ, which is why the acceptance criteria for this QC were raised to 30%. The higher variability was nevertheless insignificant, as the acceptance criteria were easily met.

On the basis of a 95% detection confidence,[12,24] the LOD was estimated to be 2.5 copies/reaction (Figure 3A), which falls within the estimated range of 2–20 determined during the development phase. Dilution linearity was demonstrated across 2 orders of magnitude for both genes and with high accuracy (Figures 3B and 3C). The hSmad7 assay was especially accurate with no more than ±2% bias at all dilutions and although accuracy was somewhat compromised with the mTBP assay at 39.9 copies/reaction (third dilution), the % bias was still well within the acceptance criteria of ±25%. These data together indicate that samples containing high amounts of hSmad7 target DNA, concentrations outside the sensitivity range, can be diluted to within the range while having a negligible and insignificant effect on accuracy. Matrix effects were also determined to be negligible as the acceptance criteria for recovery and selectivity were met for all tissues (Figure 3D). Recovery ranged from 70% to 85% for all but liver and although spleen and spinal cord samples initially failed, this was due to poor droplet generation and they passed after being diluted 10-fold and re-assayed. Matrix effects were even less pronounced on selectivity as % bias ranged from −0.25% to −11.75% for all tissues except whole blood and ovaries. This is particularly noteworthy as only 80 copies/reaction were used in these assays, suggesting that high assay sensitivity is also preserved in most if not all tissues.

Assay specificity was demonstrated in multiple ways and was assured theoretically by BLAST analysis of primer and probe sequences that detected only the intended targets. Empirical tests for specificity included non-spiked samples from the selectivity experiment and the no template control (NTC) samples, all of which were negative (data not shown). Additional evidence of specificity was demonstrated by plotting the individual fluorescent signals from the precision and accuracy experiments as only two distinct populations were detected at all QC levels, one negative and one positive for both genes (Figures 4A and 4B). Moreover, positive signal amplitudes for both genes remained constant at all QCs and there was no evidence of "rain" or intermediate amplification, characteristics indicative of high amplification efficiency.[25] This rudimentary test also helped identify "bad wells" as indicated by the single outlier above the LLOQ positive and negative signals for both hSmad7 and mTBP. Finally, assay stability was demonstrated on samples stored at room temperature (RT), at 4°C and after repeated freeze-thaw cycles (Figure 4C). Accuracy for both genes was estimated at 95% or better using high QC (HQC) and LQC samples and at all storage

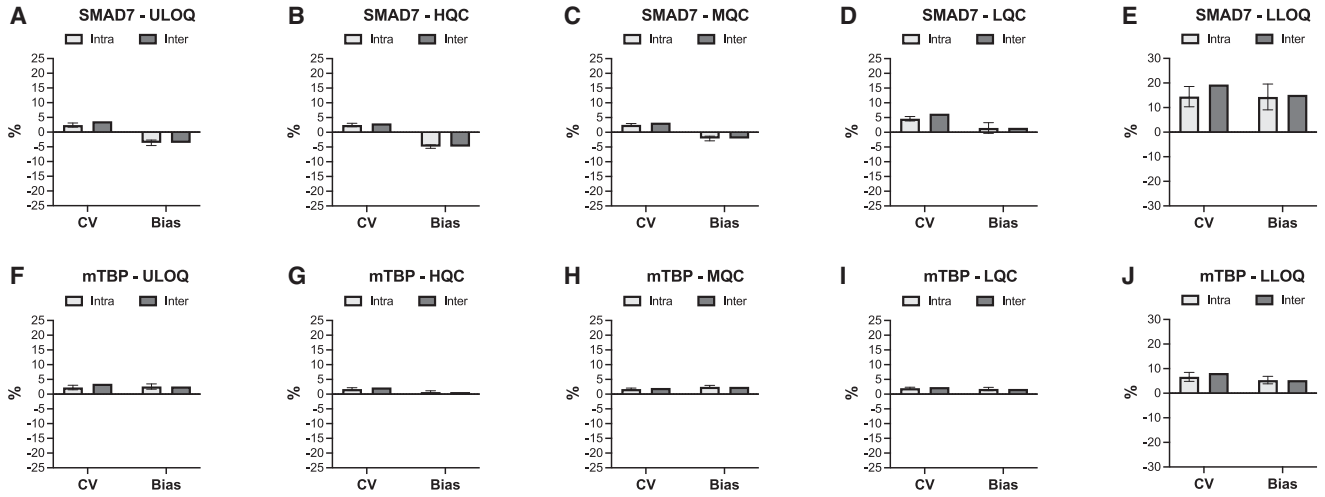


**Figure 2. Validation of precision and accuracy**

Percentage coefficient of variance (CV), a measure of precision, and bias, a measure of accuracy, were determined from amplifications at five quality control (QC) levels for hSmad7: upper limit of quantitation (ULOQ), high QC (HQC), middle QC (MQC), low QC (LQC), and lower limit of quantitation (LLOQ), respectively, corresponding to 80,000, 8,000, 800, 80, and 10 copies/reaction. Because mTBP is used for normalization, a fixed amount of mouse gDNA in each reaction (79,870 copies/reaction) was assessed at each hSmad7 QC level. Data were collected from runs 1, 5–8, and 10 (Table 3) with three plate replicates/run. Intra-assay variation was determined from means (n = 6, ±SEM) of % CV or bias run-means. Inter-assay variation was calculated from the collective mean and SD of the amplification data (i.e., copies per reaction). In each graph, y axes were bound to acceptance criteria limits.

conditions. These data together indicate that the assay is highly specific and accurate under the storage conditions assessed.

## DISCUSSION

Traditional pharmacokinetics are generally considered to be irrelevant endpoints in the safety evaluation of gene therapies. Selection of targeted clinical doses are instead guided by the results of systemic or target organ exposure to both vector and transgene. Thus, a robust analysis of vector and transgene biodistribution is a required endpoint in IND-enabling toxicology assessments for developing gene therapy programs. FDA guidance (21 CFR Part 58)[26] further stipulates that first-in-human dose determination should be conducted in compliance with GLP. Validation of developed methods is, therefore, a critical component of the process and is necessary to demonstrate robustness, linearity, precision, accuracy, and specificity of the selected assay whether it is used for biodistribution or vector titering of GMP products, which presents slightly different challenges to those described herein (e.g., distinguishing capsulated from free DNA). Despite this requirement, limited historical data are available to guide the development and validation of these key test methods. This knowledge gap is

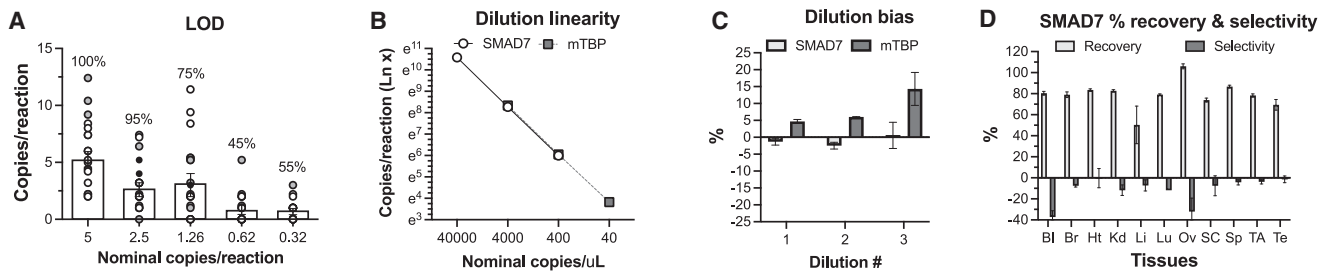


**Figure 3. Validation of detection limit, linearity, recovery, and selectivity**

(A) Plasmid DNA was serially diluted to the indicated concentrations and used in 3 separate assays for a total of 20 replicates/concentration. The shown data points are differentially shaded for each assay (white, gray, and black) while mean ± SEM values are plotted. Percentages above histograms represent the probability of positive detections. (B and C) Eight hundred thousand copies per microliter of hSmad7 plasmid DNA containing 79,870 copies/reaction of mouse gDNA were diluted to the indicated concentrations and assayed. Mean ± SEM are plotted with 3 replicates/concentration/target. Accuracy at each dilution is plotted as % bias. (D) Matrix interference on recovery and selectivity was assessed using extracts from 11 tissues (n = 2–3/tissue/assay): Bl, whole blood; Br, brain; Ht, heart; Kd, kidney; Li, liver; Lu, lungs; Ov, ovaries; SC, spinal cord; Sp, spleen; TA, tibialis muscle; Te, testis. Extractions for recovery assays were performed with 50,000 copies of spiked plasmid DNA. For selectivity, the y axis % represents bias. Assays were performed using spiked and non-spiked extracts containing 80 copies/μL added after extraction. Non-spiked samples were all negative and are not shown.

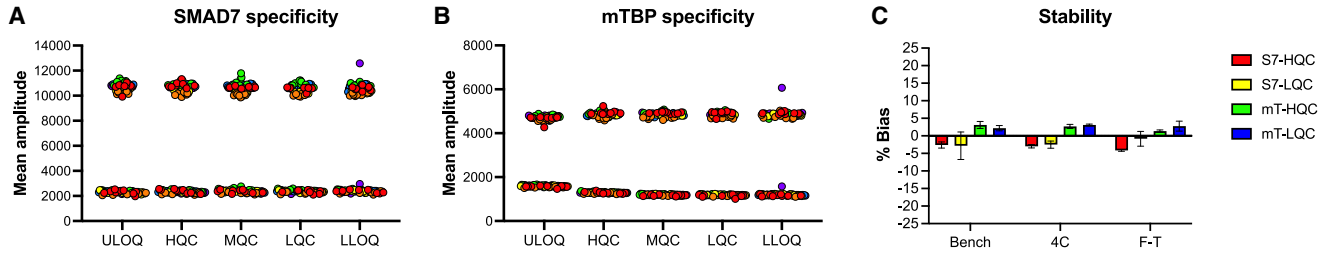


**Figure 4. Validation of specificity and stability**

(A and B) Potential amplification of non-specific targets was assessed by distinguishing unique droplet populations for positive and negative signals as multiple positive populations are indicative of non-specific amplification. Mean FAM (A) and HEX (B) amplitudes obtained from the precision and accuracy assays (Figure 3) are plotted for each well (n = 9/assay) with color coding for 6 different runs (1 color/run; n = 9/run, 54 total). Upper clusters represent positive signals for hSmad7 and mTBP, lower clusters represent negative. (C) Samples were stored at RT for 24 h (bench), at 4°C or subjected to 6 freeze-thaw cycles prior to testing (mean ± SEM shown). Assays were performed at high and low QC levels for hSmad7 and mTBP (S7- and mT-, respectively).

partly addressed by the studies reported herein that in essence provide a functional benchmark to guide future method development and validation of ddPCR assays.

The development phase assessments compared the proficiencies of different primer/probe sets in singleplex and duplex assays. These tests demonstrated primer/probe compatibility and helped determine the optimal pairing of S7.1 with mTBP2, which were then used to demonstrate accuracy and precision across a wide range of target DNA concentrations, over 4 magnitudes. Initial characterization of primer/probe sets was performed using qPCR. Although unnecessary, this proved to be valuable as visualizing the amplification curves helped identify potential problems and thus, to decide upon final primer/probe sets.

The criteria used in validating the assay were consistent with GCC recommendations and in some instances exceeded them.[12] For example, the GCC recommended bias and CV for validating accuracy and precision is ±35% and $\leq$40%, respectively, for all QCs.[12] We used ±30% and $\leq$30% for LLOQ (10 copies) and ±25% and $\leq$25% for all other QCs. The GCC also recommends assessing selectivity and sensitivity by assuring that all non-tissue control values fall below the LOD, which was the case with our assay. However, we additionally evaluated the parameters separately using a single droplet population as an indicator of specificity and ±20% bias at LQC for selectivity. It should be noted that GCC recommendations and our acceptance criteria both required a minimum of 10,000 droplets/reaction, although we averaged more than 22,000 in the validation experiments. This likely contributed to the high assay precision as this strongly correlates with droplet number.[27]

The GCC recommendations were based upon surveys of 44 different CROs of which 89% used either qPCR or ddPCR for bioanalysis. Although none of the CROs use ddPCR exclusively, 71% used both technologies for exploratory (10%) and regulated bioanalysis (90%). This includes assays for biodistribution, vector titering, vector shedding, and other aspects of gene therapy programs. Their goals were to (1) provide unified recommendations for validation, (2) facilitate interactions between sponsors and regulators, and (3) develop industry-wide internal SOPs for qPCR and ddPCR assays. Their contribution is particularly valuable as regulatory guidance for qPCR is limited and is nonexistent for ddPCR. Recent literature reviews[5,28,29] were used by the GCC to construct a consensus for qPCR, but not ddPCR, as the reviews focused primarily on qPCR often citing the inaccurate belief that ddPCR is infrequently used for regulated bioanalysis. The GCC recommendations for ddPCR are, therefore, critically important as both academic and regulatory guidance are collectively lacking. This also underscores the importance of studies like our own, which we believe is the first to be based on GCC recommendations and is therefore a good example for other developing cell and gene therapy programs.

At the time of manuscript submission, only two publications had reported the development of a ddPCR assay for gene therapy biodistribution assessments.[30,31] This likely explains the misconception that the technology is not readily used or is at least an emerging technology in the field of bioanalysis despite the GCC survey clearly indicating ddPCR is well established in the biotech industry. Kondratov et al.[31] developed ddPCR assays to quantify and track the viral load of 29 different capsids in 15 brain areas. None of the assays underwent a qualification or validation process and they were originally developed to compare AAV manufacturing efficiency in HEK293 versus Sf9 cells.[32] Sugimoto et al.[30] developed a qualified ddPCR assay for assessing CAR-T cell biodistribution and cellular kinetics using similar equipment and reagents reported herein. The study was performed prior to the publishing of GCC recommendations so assay performance used a fit-for-purpose design because of the lack of regulatory guidance and an acceptance criteria consensus. Although assay development was not described, the qualification process included accuracy and precision measures of four QC levels (HQC, middle QC [MQC], LQC, and LLOQ), LOD determination, a dilution linearity assessment and recovery and selectivity from blood extracts. Comparing this assay to our own is difficult as it was designed to track cells instead of viral vectors. Nevertheless, assay accuracy and precision were within the GCC recommendations as the inter- and intra-assay % bias range was −5.1 to −9.8, and the % CV range was 2.8 to 22.5, suggesting that ddPCR as a technology is sufficiently flexible to be used with both viral and cell-based gene therapy programs.

**Table 2. Validation parameters**

| Parameter | Definition | Method/indicator |
|---|---|---|
| Accuracy | agreement between measured and reference values | % bias at each QC level |
| Dilution linearity | measure of assay proportionality with analyte concentration | dilute spiked sample into the assay range |
| Precision | inter- and intra-assay variability with the same and different operators, respectively | % CV at each QC level |
| Quality controls (QCs) | ULOQ, upper limit of quantification<br>HQC, high QC<br>MQC, middle QC<br>LQC, low QC<br>LLOQ, lower limit of quantification<br>LOD, limit of detection | 80,000 copies/reaction<br>8,000 copies/reaction<br>800 copies/reaction<br>80 copies/reaction<br>10 copies/reaction<br>2.5 copies/reaction |
| Recovery | analyte extraction efficiency by spiking known amount of plasmid prior to DNA extraction | % of spiked concentration |
| Ruggedness/ robustness | reliability with different reagents and operators | % CV and % bias |
| Selectivity | accuracy of assay in different tissue matrices | % bias at LQC |
| Specificity | detecting only the target analyte, at different QCs and in a tissue matrix | number of droplet populations detected |
| Sensitivity | LOD, limit of detection; lowest analyte concentration detected with acceptable accuracy and precision | analyte concentration with 95% detection confidence |
| Stability | analyte detection under different storage conditions for a given amount of time | % bias with each condition/ time point |

Although our assay supports the development of AVGN7, an AAV-based gene therapeutic, the validation road map described is agnostic to therapeutic modality and could be used, for example, with other viral and even non-viral gene therapy programs. Each must overcome similar regulatory challenges with limited agency guidance, particularly the latter, which illustrates the broad significance and potential impact of our studies. Indeed, the strict adherence to rigorous standards[12] set by technical populists in the field and refined with subsequent studies such as ours will likely interest regulatory authorities and even influence guidance decisions and documents. This in turn will provide stability to what is arguably the most dynamic sector of drug development, gene therapy.

## MATERIALS AND METHODS

### Overview

The ddPCR assay was constructed in two phases: development and validation. The former optimized conditions and demonstrated theoretical limits, ranges, and sensitivities using customized analyte and statistical analyses. By contrast, the latter used empirical data to demonstrate assay reliability and reproducibility under conditions of intended use. The terms/metrics used in each phase, therefore, are not necessarily interchangeable and are defined below.

Initial studies characterized the general amplification characteristics of 4 different primer/probe sets, 2 each for the codon-optimized hSmad7 expression construct contained within the AVGN7 viral genome as well as the mTBP, a reference gene commonly used for total DNA normalization in PCR applications.[33,34] Assay performance was then tested in singleplex and duplex reactions to determine compatibility of the 6 potentially interacting oligonucleotides before deciding upon a final combination of primers and probes to target each template. Assay validation was performed under GLP compliance and evaluated the following parameters: accuracy/precision, ruggedness/robustness, selectivity, recovery, specificity, dilution linearity, and target gene stability (Table 2). These data were collectively obtained from a number of experiments performed by different operators, on different dates using different reagent stocks, all of which are required for assay validation (Table 3). The procedures described below include those used in both assay development and validation as well as the specific experiments required of each phase.

### Critical reagents and equipment

Droplet Generation Oil, ddPCR Droplet Reader Oil, and 2X ddPCR SuperMix for Probes (no dUTP) were purchased from Bio-Rad. Primers and probes (see below) were purchased from Integrative DNA Technologies as was gBlock reference DNA. The latter was double-stranded DNA corresponding to the predicted amplicon sequences and was stored at concentrations of 1e9 copies/μL in 1:1 Tris (10 mM)-EDTA (1 mM) (pH 7.8)/glycerol with 1 μg/mL sonicated salmon sperm gDNA (Abnova). The transfer plasmid used in manufacturing AVGN7 was also used as a reference standard and was generated by GenScript. All ddPCR assays were performed using the Bio-Rad system. This includes the QX200 Droplet Generator, the QX200 Droplet Reader, the C1000 Touch PCR Thermal Cycler, the PX1 Plate Sealer, and various consumables (e.g., plates, heat seal foil, droplet cartridges) designed to work with the system. The Quantstudio 7 Pro qPCR System (Applied Biosystems) was also used in the initial assessment of primer/probe sets.

### Primer and probe design

Two primer and probe sets for each gene target (Table 4) were designed to operate with the Bio-Rad QX200 ddPCR system. The codon-optimized human Smad7 (hSmad7) and mTBP target gene sequences were first prescreened for secondary structure. Primers were then chosen on the basis of melting temperature ($T_m$) (55-65°C), amplicon size (75–200 bp), GC content (50%–60%), and the absence of primer predicted secondary structure in the presence of 50 mM salt and 300–900 nM oligonucleotide concentrations. We initially neglected to verify the gDNA annealing sites and later learned that the mTBP1 primers mapped to exons 2 and 3. Although this produces a 1,280 bp amplicon, the large amplicon size did not ultimately affect assay performance. Probes were designed to have $T_m$ values 5°C–10°C above the primers. Carboxyfluorescein (FAM) and

**Table 3. Summary of ddPCR validation schedule**

| Run/day | Operator | Purpose |
|---|---|---|
| 1 | a | accuracy/precision and limit of detection |
| 2 | a | accuracy/precision and limit of detection |
| 3 | a | accuracy/precision and limit of detection |
| 4 | a | accuracy/precision and dilution linearity |
| 5 | b | accuracy/precision and selectivity |
| 6 | b | selectivity and recovery |
| 7 | a | accuracy/precision and stability (benchtop and refrigerator) |
| 8 | a | selectivity |
| 9 | b | selectivity and stability (freeze-thaw) |
| 10 | b | selectivity and recovery |
| 11 | b | selectivity and recovery |

Runs performed sequentially on different days. Operator letters correspond to individuals performing stated task.

hexachlorofluorescein (HEX) dyes were separately paired with Black Hole quenchers on the hSmad7 and mTBP probes, respectively, as per Bio-Rad recommendations. Because the hSmad7 payload gene sequence is codon optimized, primer and probe sequences were incompatible with mouse Smad7 DNA. Nevertheless, BLAST analysis was used to confirm specificity of all primers and probes.

### Tissue processing

Genomic DNA was extracted from whole mouse blood and various other mouse tissues to assess matrix effects, recovery, and selectivity. Whole blood and frozen tissues (tibialis muscle, liver, brain, spleen, heart, lung, spinal cord, kidney, ovary, and testes) from C57BL/6 mice were purchased from BioIVT as no animals were used in these studies. The procedures described are identical to those used when preparing samples for biodistribution assays under GLP compliance and employ the KingFisher Flex Processor (Thermo Fisher Scientific), an automated nucleic acid purification system, combined with the MagMAX DNA Multiple Sample Ultra 2.0 reagent kit (Applied Biosystems). Tissue samples were first homogenized using a Precellys 24 homogenizer and CK28 lysing kit followed by overnight Proteinase K digestion. Blood samples were not homogenized and were treated with Proteinase K just prior to loading into the KingFisher, but were otherwise similarly processed. DNA quantification was performed using the Quant-iT PicoGreen kit (Thermo Fisher Scientific), in 96-well format and at RT using a DNA standard curve range of 62.5–4,000 ng/mL. Manufacturers' protocols were used for each kit and system.

### Assay development

Initial assessments of singleplex primer/probe sets were conducted using qPCR to verify annealing temperatures and to visualize amplification curves. gBlock concentrations from 0 to 1e4 copies/μL were amplified for 45 cycles using 900 nM primers, 250 nM probes and the following amplification program: 95°C for 10 min, 40 cycles of 95°C for 15 s, then 60°C for 1 min with a ramp rate of 1.6°C/s between cycles. All other experiments were performed using ddPCR and the following amplification program: 95°C for 10 min, 40 cycles of 95°C for 30 s, then 59°C for 1 min, followed by 98°C for 10 min with a ramp rate of 2°C/s between cycles. Optimal concentrations of primers and probes used for ddPCR were determined by titering both with a 30% dilution scheme for probes (85.75, 122.5, 175, and 250 nM) and a 50% dilution scheme for primers (112.5, 225, 450, and 900 nM). This was performed independently for each primer/probe set (Table 4) in singleplex reactions, although with 1,000 copies/μL gBlock for mTBP sets and 10 copies/μL for hSmad7 sets. The different amounts of template were used to emulate the expected differences when performing assays on tissue extracts. Optimal primer/probe concentrations were then used to amplify 8, 80, 800, 8,000, and 80,000 copies/reaction in singleplex reactions for each target gene. Compatibility of hSmad7 primer/probe sets with mTBP sets was determined with duplex reactions using the same template concentration scheme. Once the choice of primer/probe sets were finalized (hSmad7.1 and mTBP2), preliminary estimates of assay range, linearity, precision, accuracy, sensitivity (LOD) and LLOQ (50 copies/μg DNA[35,36]) were evaluated by amplifying 2, 20, 200, 2,000, 40,000, and 80,000 copies/reaction of gBlock template (n = 6/concentration).

### Assay validation

Prior to initiating these studies, a validation plan was generated to describe SOPs for all laboratory activities as well as a system to assure regulatory compliance that included quality assurance, data collection, and record retention. Because the FDA has yet to produce guidance documents for PCR-based assays, especially ddPCR, we consulted guidance documents for bioanalytical method validation,[2] nonclinical biodistribution considerations[37] and preclinical assessments[1] as well as relevant reviews.[5,11] Each amplification plate included an NTC sample of nuclease-free water to ensure noncontamination. Validation QCs (Table 2) were prepared by spiking the linearized transfer plasmid used to manufacture AVGN7 (phSmad7) into a matrix that contains mouse genomic DNA (mgDNA) (a total of 250 ng DNA/reaction or 79,870 copies of mTBP/reaction on the basis of theoretical mTBP gene copy number) and stored at −20°C. All QCs were analyzed in the accuracy/precision experiments, only LQC and HQC in non-accuracy/precision experiments, and no QCs were required for the LOD experiments.

Accuracy/precision was established collectively in 6 separate experiments performed by two different operators (Table 3) across multiple days, which assesses ruggedness/robustness. In addition, two different master mix lots were used. Triplicate samples were run for each QC and each sample was run in triplicate. Sample means were then used to calculate mean values for each QC (n = 3/plate/experiment). Limits of quantification were defined by ULOQ and LLOQ and inter- and intra-assay results were assessed according to a target acceptance criteria (see below). Two-thirds (66.7%) of the accuracy/precision runs were required to meet these criteria. If necessary, accommodations for failures were developed to include outlier tests (boxplot 1.5 × IQR or 3 × IQR) on the cumulative dataset.

**Table 4. Primers and probes**

| Assay | Primer/probe | Sequence 5′–3′ | $T_m$ | Amp |
|---|---|---|---|---|
| hSmad7.1 | S7.1-F<br>S7.1-R<br>S7.1-P | confidential intellectual property; probe labeled with FAM and BHQ1 | 59<br>58<br>65 | 111 |
| hSmad7.2 | S7.2-F<br>S7.2-R<br>S7.2-P | confidential intellectual property; probe labeled with FAM and BHQ1 | 57<br>57<br>65 | 135 |
| mTBP1 | T1-F<br>T1-R<br>T1-P | GAAGAAAGGGAGAATCATGGACGAGTAA<br>GTCCTGTGCCGTAAGHEX - CCTGAGCATA<br>AGGTGGAAGGCTGTT - BHQ1 | 55<br>57<br>62 | 1,280 |
| mTBP2 | T2-F<br>T2-R<br>T2-P | ACCCCACAACTCTTCCATTCGGGTCATAG<br>GAGTCATTGGTGHEX - ACGGTGCAGTGGT<br>CAGAGTTTGAG - BHQ1 | 54<br>56<br>62 | 91 |

$T_m$, melting temperature (°C); Amp, amplicon base pairs; F/R, forward and reverse primer; P, probe; FAM, carboxyfluorescein dye; HEX, hexachlorofluorescein dye; BHQ1, black hole quencher.

LOD was defined as the lowest copy number that the assay can detect 95% of the time. This is consistent with ddPCR recommendations from the GCC and from the European Network of GMO Laboratories[12,24] but differs from quantitative qPCR recommendations developed by the GCC[12] and others[5,38] as well as from qualitative qPCR recommendations.[39] This is an important distinction as unlike ddPCR, qPCR assays rely upon standard curves or are strictly qualitative. Five independent preparations of serially diluted phSmad7 samples were used to determine LOD: 5, 2.5, 1.25, 0.625, and 0.313 copies/reaction. Samples were diluted in a matrix containing mouse gDNA and each dilution was evaluated with 20 determinations (n = 20) across multiple runs. LOD for the hSmad7 assay was defined as the lowest concentration providing 95% detection confidence (i.e., 95.0% of samples indicate >0 copies/reaction). Dilution linearity was established by diluting phSmad7 from a concentration above ULOQ to within the assay range. The initial intermediate sample contained 40,000 copies of hSmad7 in 31.25 ng/μL mouse gDNA. It was then diluted serially in 31.25 ng/μL single-stranded DNA (ssDNA) resulting in 40,000, 4,000, and 400 copies/reaction of phSmad7 and 3,994, 399.4 and 39.9 copies/μL of mTBP, target gene concentrations above and within the detection range. Three replicates of each dilution level were evaluated in each experiment using n = 3/sample.

Selectivity was assessed using various tissue extracts and whole blood from two naive C57BL/6 mice. Non-spiked extracts and those containing 250 ng (80.0 copies/reaction) of phSmad7 were evaluated with 3 tissue replicates. Non-spiked samples were expected to have values below LLOQ, while 80% of the phSmad7 spiked samples were expected to have % bias values within ±25.0 of the nominal concentration. For mTBP, 80% of both non-spiked and spiked samples were expected to have % bias values within ±25.0 of 79,870 copies/reaction. For recovery experiments, tissues were divided and 50,000 copies of phSmad7 were added to one-half before extracting. Both spiked and non-spiked samples were run undiluted and % recovery was calculated as a ratio to the expected 100%. The high concentration total DNA (i.e., gDNA plus spiked phSmad7) in the undiluted sample presents the opportunity of reagent depletion. Thus, only hSmad7 primers and probes were used in singleplex reactions and each sample was run in triplicate. At least 30% of the spiked concentration was expected to be recovered for at least one of the tissue replicates.

Specificity was demonstrated by verifying that the hSmad7 cDNA concentrations in all selectivity naive samples (non-spiked) and in NTC controls were below LOD. It was also assessed by visualizing signal separation between positive and negative droplet populations in the accuracy and precision experiments. The fluorescence amplitude for all droplets in the FAM and HEX channels were plotted after threshold calculation using QuantaSoft where each droplet was represented by a single dot in the 2-D Amplitude view. Multiple positive populations are indicative of non-specific amplifications whereas a single positive population indicates high specificity. Stability was tested using HQC and LQC single-use aliquots stored at −20°C, and three samples of each QC were tested in triplicate for each stability time point. Long-term stability was not assessed in light of the very well documented stability of DNA in 10 mM Tris, 1 mM EDTA (pH 8.0) at −20°C.[5] Freeze-thaw stability was evaluated after six freeze-thaw cycles on 12 h intervals of freezing to −20°C and thawing to RT for 30 min. Refrigerator and benchtop stability was evaluated after storing thawed samples at 2°C–8°C or at RT, respectively, for at least 24 h.

### General validation acceptance criteria

The criteria used were more stringent than the recently established consensus on validation parameters and with the relevant procedures.[12] These criteria were initially established for the accuracy/precision evaluation and were based on the percent bias for accuracy and percent CV for precision. These values were calculated from the means of triplicate measures for each QC level; 30% was used for LLOQ and 25% for other levels. If percentages exceeded these levels, a mechanism for outlier removal was applied where each replicate value would be individually removed, allowing the recalculation of a new value using the remaining duplicates. Failure occurs when all iterations for a particular QC level exceed the criteria and would result

in the experiment being repeated. After the accuracy/precision criteria were met, the criteria were revised and adopted for use in all subsequent validation runs.

Tests of parameters other than accuracy/precision or to determine LOD included at least one set of QCs (LQC and HQC) and an NTC, all run in triplicate. Acceptance for LQC and HQC was set at 25% CV and bias. A sample was considered negative when at least two of the three replicates had less than 4 positive droplets. The generated droplet acceptance criterium was set at 10,000/well. Any well with fewer droplets was excluded from analysis and at least 2 wells were required to meet this criterium to define a final concentration. Data interpretation was contingent upon run acceptance for LQC, HQC, and NTC. If any of these failed, the validation experiment(s) would be repeated.

### Data analysis

The CV (SD/mean) for any given group was calculated as a measure of precision and is expressed as a percentage. Similarly, the measure of accuracy or bias was also expressed as a percentage and is the difference between nominal and observed copies. These values were calculated from group means, not from replicates of each sample.

## DATA AVAILABILITY STATEMENT

All non-proprietary data are available upon request to the corresponding author.

## ACKNOWLEDGMENTS

These studies were supported by grants from the National Institutes of Health to B.D.R. (R44CA221539 and R44AG079825). The studies reported herein were performed in Rockville, MD and Ewing, NJ, USA.

## AUTHOR CONTRIBUTIONS

B.D.R., conceptualization, formal analysis, funding acquisition, resources, writing (original draft, review and editing), supervision, and visualization; S.K.H., conceptualization, project administration, writing (review and editing); D.R.C. and J.C., investigation and writing (review and editing); A.G.R., data curation, formal analysis, methodology, project administration, writing (review and editing), supervision, and validation.

## DECLARATION OF INTERESTS

B.D.R. is the founder, CEO, and majority stock owner of AAVogen, Inc., which develops the proprietary gene therapeutic for muscle wasting (AVGN7) that is discussed herein. S.K.H. directs AAVogen's preclinical development and regulatory affairs, while the other authors are staff scientists of Smithers. Neither S.K.H. nor the other authors have any financial interests to disclose.

## REFERENCES

1. Center for Biologics Evaluation and Research (2013). Preclinical assessment of investigational cellular and gene therapy products: guidance for industry. In US Department of Health and Human Services, FaDA.
2. Center for Drug Evaluation and Research, and Center for Veterinary Medicine (2018). Bioanalytical method validation: guidance for industry. In US Department of Health and Human Services, FaDA.
3. Pharmaceuticals and Medical Devices Agency (2019). Japan Guidance on Ensuing the Quality and Safety of Products for Gene Therapy.
4. European Medicines Agency (2009). General Principles to Address Virus and Vector Shedding.
5. Ma, H., Bell, K.N., and Loker, R.N. (2021). qPCR and qRT-PCR analysis: regulatory points to consider when conducting biodistribution and vector shedding studies. Mol. Ther. Methods Clin. Dev. *20*, 152–168.
6. Basu, A.S. (2017). Digital assays Part I: partitioning statistics and digital PCR. SLAS Technol. *22*, 369–386.
7. Majumdar, N., Wessel, T., and Marks, J. (2015). Digital PCR modeling for maximal sensitivity, dynamic range and measurement precision. PLoS One *10*, e0118833.
8. Ishak, A., AlRawashdeh, M.M., Esagian, S.M., and Nikas, I.P. (2021). Diagnostic, prognostic, and therapeutic value of droplet digital PCR (ddPCR) in COVID-19 patients: a systematic review. J. Clin. Med. *10*, 5712.
9. Li, H., Bai, R., Zhao, Z., Tao, L., Ma, M., Ji, Z., Jian, M., Ding, Z., Dai, X., Bao, F., and Liu, A. (2018). Application of droplet digital PCR to detect the pathogens of infectious diseases. Biosci. Rep. *38*.
10. Lock, M., Alvira, M.R., Chen, S.J., and Wilson, J.M. (2014). Absolute determination of single-stranded and self-complementary adeno-associated viral vector genome titers by droplet digital PCR. Hum. Gene Ther. Methods *25*, 115–125.
11. Gorovits, B., Marshall, J.C., Smith, J., Whiteley, L.O., and Neubert, H. (2019). Bioanalysis of adeno-associated virus gene therapy therapeutics: regulatory expectations. Bioanalysis *11*, 2011–2024.
12. Wissel, M., Poirier, M., Satterwhite, C., Lin, J., Islam, R., Zimmer, J., Khadang, A., Zemo, J., Lester, T., Fjording, M., et al. (2022). Recommendations on qPCR/ddPCR assay validation by GCC. Bioanalysis *14*, 853–863.
13. Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. *55*, 611–622.
14. dMIQE Group, and Huggett, J.F. (2020). The digital MIQE guidelines update: minimum information for publication of quantitative digital PCR experiments for 2020. Clin. Chem. *66*, 1012–1029.
15. Rodgers, B.D., and Ward, C.W. (2022). Myostatin/activin receptor ligands in muscle and the development status of attenuating drugs. Endocr. Rev. *43*, 329–365.
16. Maricelli, J.W., Bishaw, Y.M., Wang, B., Du, M., and Rodgers, B.D. (2018). Systemic SMAD7 gene therapy increases striated muscle mass and enhances exercise capacity in a dose-dependent manner. Hum. Gene Ther. *29*, 390–399.
17. Winbanks, C.E., Murphy, K.T., Bernardo, B.C., Qian, H., Liu, Y., Sepulveda, P.V., Beyer, C., Hagg, A., Thomson, R.E., Chen, J.L., et al. (2016). Smad7 gene delivery prevents muscle wasting associated with cancer cachexia in mice. Sci. Transl. Med. *8*, 348ra98.
18. Himeda, C.L., Chen, X., and Hauschka, S.D. (2011). Design and testing of regulatory cassettes for optimal activity in skeletal and cardiac muscles. Methods Mol. Biol. *709*, 3–19.
19. Gonçalves, M.A., Janssen, J.M., Nguyen, Q.G., Athanasopoulos, T., Hauschka, S.D., Dickson, G., and de Vries, A.A.F. (2011). Transcription factor rational design improves directed differentiation of human mesenchymal stem cells into skeletal myocytes. Mol. Ther. *19*, 1331–1341.
20. Salva, M.Z., Himeda, C.L., Tai, P.W., Nishiuchi, E., Gregorevic, P., Allen, J.M., Finn, E.E., Nguyen, Q.G., Blankinship, M.J., Meuse, L., et al. (2007). Design of tissue-specific regulatory cassettes for high-level rAAV-mediated expression in skeletal and cardiac muscle. Mol. Ther. *15*, 320–329.
21. Mullard, A. (2021). Gene therapy community grapples with toxicity issues, as pipeline matures. Nat. Rev. Drug Discov. *20*, 804–805.
22. Morales, L., Gambhir, Y., Bennett, J., and Stedman, H.H. (2020). Broader implications of progressive liver dysfunction and lethal sepsis in two boys following systemic high-dose AAV. Mol. Ther. *28*, 1753–1755.

23. Center for Biologics Evaluation and Research (2021). Food and Drug Administration Cellular, Tissue, and Gene Therapies Advisory Committee (CTGTAC) Meeting #70; Toxicity Risks of Adeno-Associated Virus (AAV) Vectors for Gene Therapy (GT). In: Administration, H.HSF.D (ed). Washington, D.C. p 69.

24. Pecoraro, S., Berben, G., Burns, M., Corbisier, P., De Giacomo, M., De Loose, M., Dagand, E., Dobnik, D., Eriksson, R., Holst-Jensen, A., et al. (2019). Overview and Recommendations for the Application of Digital PCR.

25. Lievens, A., Jacchia, S., Kagkli, D., Savini, C., and Querci, M. (2016). Measuring digital PCR quality: performance parameters and their optimization. PLoS One *11*, e0153317.

26. Archives, N.. CoFR Good Laboratory Practice for Nonclinical Laboratory Studies. www.ecfr.gov/current/title-21.

27. Pinheiro, L., and Emslie, K.R. (2018). Basic concepts and validation of digital PCR measurements. Methods Mol. Biol. *1768*, 11–24.

28. Yang, T.Y., and Doddareddy, R. (2021). Considerations in the development and validation of real-time quantitative polymerase chain reaction and its application in regulated bioanalysis to characterize the cellular kinetics of CAR-T products in clinical studies. Bioanalysis *13*, 115–128.

29. Corsaro, B., Yang, T.Y., Murphy, R., Sonderegger, I., Exley, A., Bertholet, S., Dakappagari, N., Dessy, F., Garofolo, F., Kierstead, L., et al. (2021). 2020 white paper on recent issues in bioanalysis: vaccine assay validation, qPCR assay validation, QC for CAR-T flow cytometry, NAb assay harmonization and ELISpot validation (Part 3 - recommendations on immunogenicity assay strategies, NAb assays, biosimilars and FDA/EMA immunogenicity guidance/guideline, gene & cell therapy and vaccine assays). Bioanalysis *13*, 415–463.

30. Sugimoto, H., Chen, S., Minembe, J.P., Chouitar, J., He, X., Wang, H., Fang, X., and Qian, M.G. (2021). Insights on droplet digital PCR-based cellular kinetics and biodistribution assay support for CAR-T cell therapy. AAPS J. *23*, 36.

31. Kondratov, O., Kondratova, L., Mandel, R.J., Coleman, K., Savage, M.A., Gray-Edwards, H.L., Ness, T.J., Rodriguez-Lebron, E., Bell, R.D., Rabinowitz, J., et al. (2021). A comprehensive study of a 29-capsid AAV library in a non-human primate central nervous system. Mol. Ther. *29*, 2806–2820.

32. Kondratov, O., Marsic, D., Crosson, S.M., Mendez-Gomez, H.R., Moskalenko, O., Mietzsch, M., Heilbronn, R., Allison, J.R., Green, K.B., Agbandje-McKenna, M., et al. (2017). Direct head-to-head evaluation of recombinant adeno-associated viral vectors manufactured in human versus insect cells. Mol. Ther. *25*, 2661–2675.

33. Kumar, S., Ahmad, A., Kushwaha, N., Shokeen, N., Negi, S., Gautam, K., Singh, A., Tiwari, P., Garg, R., Agarwal, R., et al. (2022). Selection of ideal reference genes for gene expression analysis in COVID-19 and mucormycosis. Microbiol. Spectr. *10*, e0165622.

34. Nakao, R., Yamamoto, S., Yasumoto, Y., Kadota, K., and Oishi, K. (2015). Impact of denervation-induced muscle atrophy on housekeeping gene expression in mice. Muscle Nerve *51*, 276–281.

35. Center for Drug Evaluation and Research, and Center for Biologics Evaluation and Research (2020). Long term follow-up after administration of human gene therapy products: guidance for industry. In US Department of Health and Human Services, FaDA.

36. Center for Biologics Evaluation and Research (2006). Gene therapy clinical trials: guidance for industry. In US Department of Health and Human Services, FaDA.

37. Center for Biologics Evaluation and Research (2021). Food and drug administration (C.f.B.E.a.R.) S12 nonclinical biodistribution considerations for gene therapy products. In US Department of Health and Human Services, FaDA.

38. Forootan, A., Sjöback, R., Björkman, J., Sjögreen, B., Linz, L., and Kubista, M. (2017). Methods to determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR). Biomol. Detect. Quantif. *12*, 1–6.

39. Broeders, S., Huber, I., Grohmann, L., Berben, G., Taverniers, I., Mazzara, M., Roosens, N., and Morisset, D. (2014). Guidelines for validation of qualitative real-time PCR methods. Trends Food Sci. Technol. *37*, 115–126.