Molecular Therapy
Methods & Clinical Development
Review

PCR-based analytics of gene therapies using adeno-associated virus vectors: Considerations for cGMP method development

Emmanuel Blay,[1] Elaine Hardyman,[1] and Wesley Morovic[1]

[1]Gene & Cell Therapy, PPD GMP Laboratories, Part of ThermoFisher Scientific, Middleton, WI, USA

The field of gene therapy has evolved and improved so that today the treatment of thousands of genetic diseases is now possible. An integral aspect of the drug development process is generating analytical methods to be used throughout clinical and commercial manufacturing. Enumeration and identification assays using genetic testing are critical to ensure the safety, efficacy, and stability of many active pharmaceutical ingredients. While nucleic acid-based methods are already reliable and rapid, there are unique biological, technological, and regulatory aspects in gene therapies that must be considered. This review surveys aspects of method development and validation using nucleic acid-based testing of gene therapies by focusing on adeno-associated virus (AAV) vectors and their co-transfection factors. Key differences between quantitative PCR and droplet digital technologies are discussed to show how improvements can be made while still adhering to regulatory guidance. Example validation parameters for AAV genome titers are described to demonstrate the scope of analytical development. Finally, several areas for improving analytical testing are presented to inspire future innovation, including next-generation sequencing and artificial intelligence. Reviewing the broad characteristics of gene therapy assessment serves as an introduction for new researchers, while clarifying processes for professionals already involved in pharmaceutical manufacturing.

INTRODUCTION

Whereas a few decades ago the development of gene therapies was facing major difficulties because of adverse effects in early clinical trials, the recent regulatory approvals of therapies like voretigene neparvovec (Luxturna), onasemnogene abeparvovec (Zolgensma), and delandistrogene moxeparvovec (ELEVIDYS) foreshadow the treatments and possible cures for the estimated 5,600 diseases caused by defective genes.[1] As of April 2023, 24 gene therapies have been approved for clinical use with almost 2,022 gene therapies in development.[2] The U.S. Food and Drug Administration (FDA) defines human gene therapy as a process to "modify or manipulate the expression of a gene or to alter the biological properties of living cells for therapeutic use."[3] Advances in genome editing strategies and their delivery by vectors, especially engineered viruses, bacteria, and lipid nanoparticles, offer precise targeting, increased efficacy, broad tropism, and reduced side effects.[4–6] The FDA and other regulatory agencies provide many guidance documents for the various phases of gene therapy drug development, including good laboratory practices for pre-clinical work,[7] good clinical practices used during clinical investigations,[8] and good manufacturing practices (GMP) to produce active pharmaceutical ingredients (APIs) before treatment in humans or animals.[9] The FDA recommends that well-established, safe, and effective product delivery methods as well as accurate dosage and dose schedules are required to avoid any adverse effect in preclinical and clinical trials of vector-based therapy.[10] Previous clinical programs have shown that there is a positive correlation between the presence of gene copy numbers and protein expression.[11,12] Conversely, there are reports of treatment-emergent serious adverse events that occur after treatment has started in gene therapy studies.[13] Dose-dependent adverse effects such as hepatotoxicities, thrombotic microangiopathies, neuropathological findings, and in some cases death of the subject have been reported during treatment with adeno-associated virus (AAV).[14] Precise, accurate, robust, and reliable quantification of gene therapy preparations is important for ensuring therapeutic efficacy and safety during preclinical, clinical, and market use.[15]

One of the most common tools for testing nucleic acids is PCR. PCR-based genetic analysis is also a core technique used to assess various attributes of gene therapy manufacturing, since the API, vector, and manufacturing components all commonly have genetic attributes. Nonetheless, regulatory guidance on developing and validating PCR-based techniques is still improving in consideration of new technologies and strategies. Many reviews propose recommendations for criteria such as assay development and validation parameters and these are being adopted for use in gene therapies.[16,17] However, there remain many challenges in the various aspects of gene therapy treatment and vectors that will need to be considered in the context of these recommendations. This review surveys key biological, technological, and regulatory characteristics of nucleic acid-based analytical testing to for researchers and professionals considering future

https://doi.org/10.1016/j.omtm.2023.101132.

Correspondence: Wesley Morovic, Building 13, 3225 Deming Way, Middleton, WI 53562, USA.
E-mail: wesley.morovic@ppd.com

investigational new drugs and biologics license application filings for gene therapies. Current GMP (cGMP) methods can be defined as procedures that are developed to be fit for purpose before being validated using the most updated regulatory guidelines to generate reliable data about a target attribute of a specific drug in a locked-down approach. Biological attributes of viral vectors, particularly AAVs, are described in detail to highlight common considerations for PCR-based assay design. The technologies involved in the manufacturing of AAV vectors and genetic quality testing required during and after gene therapy manufacturing are examined. We focus on real-time qPCR and droplet digital PCR (ddPCR) technologies to highlight the benefits and challenges of both platforms when used with gene therapies. In addition to introducing key analytical technologies, specific regulatory requirements to consider for gene therapy method development are highlighted. An example cGMP quantitative assay for AAV genome titer is described to demonstrate the scope of assay validation. Last, considerations for future innovations in methodology and technology that have the potential to vastly improve cGMP testing are described.

SURVEY OF ELEMENTS TO CONSIDER FOR GENE THERAPY METHOD DEVELOPMENT

Biological characteristics of AAV that influence analytical methods

Viral vectors are one of the most popular mechanisms for delivering genetic information to a cell, thus providing the means to modify specific cell types transiently or permanently to express therapeutic genes.[18] The main viral vectors used currently are the adenoviruses (AVs), retroviruses (γ-retroviruses and lentiviruses), poxviruses, AAVs, baculoviruses, and herpes simplex viruses.[19,20] These viral vectors are broadly categorized according to whether their genomes integrate into host genome (retroviruses and lentiviruses) or persist in the cell nucleus predominantly as extrachromosomal episomes (AAV, AV, and herpes viruses).[21] The use of recombinant AAV (rAAV) has increasingly and systematically become an important vector for gene therapy and will be used as the primary vector example in this review.[22,23] For disease treatment, AAVs have become popular for therapy design because of efficiency of transgene expression, safety, selective tropism, prolonged transgenic expression, and moderate packaging capacity.[24] They also have attributes favorable in manufacturing, such as rapid onset of replication in industrial cell culture, scalability, and storage stability.[24,25] Although the biology, interactions, and therapeutic applicability of AAVs have been extensively reviewed,[26,27] several factors are important to describe here for consideration of cGMP method development.

AAVs are part of the *Dependoparvovirus* genus and composed of at least 12 known naturally occurring serotypes.[27] All AAV serotypes are small, non-enveloped viruses with single-stranded DNA (ssDNA) genomes of 4.7 kb encapsidated by an icosahedral shell.[28] The AAV genomes contain two open reading frames that encode for the non-structural replication (*rep*) and the capsid (*cap*) genes regulated by three promoters.[29,30] Two flanking inverted terminal repeats (ITRs) are non-coding, hairpin structures that provide self-priming activity for replication and packaging.[27] Because the ITRs are the only elements required for genome packaging, the *rep* and *cap* regions of the genome can be replaced by transgenes designed for therapeutic effects. An important aspect of designing genetic assays for AAVs is the secondary structure of the genome. Although their genomes are ssDNA, when delivered to the nucleus a transgene needs to be converted to double-stranded DNA (dsDNA).[31,32] Interestingly, this process may generate a dimer ssDNA with a hairpin in the middle of the two genome regions.[31] The rAAV genome structure is important to consider when designing molecular methods like PCR that require dsDNA templates and particular primer conditions to ensure efficiency, especially considering the other genetic elements used for rAAV manufacturing.

rAAV manufacturing technologies

There are several different innovative strategies to produce rAAV by transferring various genetic elements into eukaryotic cell lines to be expressed and purifying the filled capsids.[25,33] For example, Figure 1 illustrates a typical three-plasmid strategy for producing rAAV by transient transfection (TT) in human embryonic kidney (HEK) 293 cells. The plasmids encode the rAAV genome as well as exogenous sources of *rep* and *cap* to facilitate replication, capsid construction, and genome packaging.[34] Additionally, AAVs are replication deficient without plasmid-encoded helper components from other viruses or even drastic environmental stress.[27] Without these helper factors, AAVs deliver their genomes into the host cell, where they are mostly degraded after a short time, although some copies persist long term. When all the required factors are transduced, the endogenous machinery of the host cells uses the plasmids to replicate the rAAV genome and produce viral proteins. After rAAV capsid packaging, the infected host cells are lysed to release therapeutic capsids, which then can be purified using various methods outside, which are the scope of this review.

While transfection into HEK 293 cells is a common platform for rAAV production, other manufacturing strategies have unique benefits and challenges to be considered in analytical method development. Baculovirus expression vector systems (BEVSs) deliver the *rep* and *cap* functions, as well as the rAAV ITR and gene of interest (GOI) sequences, to cells of insects such as *Spodoptera frugiperda* Sf9.[35] BEVS can be scaled up in larger production volumes compared with TT; however, the components have been shown to have questionable stability.[33,36] Similarly, herpes simplex virus 1-assisted vector expression (HAVE) can be used to infect *rep*, *cap*, and helper functions into cell lines without the complexity of large-scale transfection.[37,38] A different strategy for rAAV production involves the stable engineering of the rAAV genome into producer cell lines (PCLs), such as oncogene-containing HeLa cells.[39] This allows for direct scaling and consistence performance using an AV as the helper element.[33] Recent advances in AV helper engineering have further increased the stability of PCLs and yield of rAAV while reducing cost and complexity from requiring multiple process components.[40]

Despite the production process, residual genetic elements like process plasmids and host cell DNA in TT and helper AV particles in PCL are

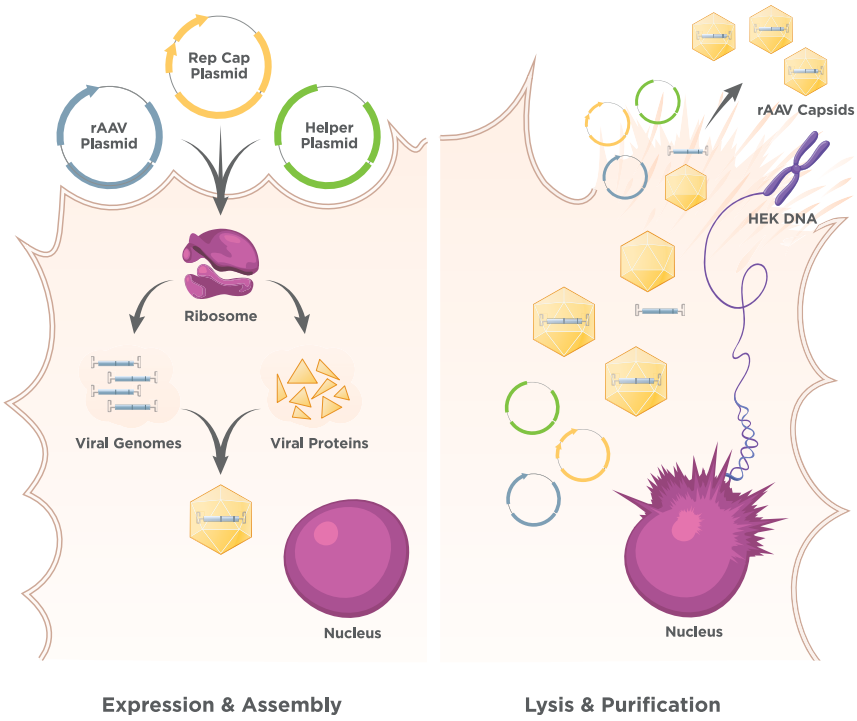

Figure 1. Summary of rAAV production and targets for analytical methods

The example three-phase manufacturing scheme for rAAV shows the different plasmids that are transfected into HEK 293 cells to produce complete rAAV capsids. (Left) The machinery of the HEK 293 cells transcribes the plasmid-encoded genes and cofactors to replicate the rAAV genome which are encapsidated by viral proteins. (Right) The manufacturing process includes a lysis step that liberates the encapsidated rAAV capsids, which are then purified for therapeutic use. All nucleic-acid sources (i.e., rAAV, RepCap, and helper plasmids; HEK 293 genomic dsDNA; rAAV genomic ssDNA) represent potential targets for cGMP analytical assessment.

impurities that could decrease efficacy, compromise product stability, introduce toxicity, and increase a recipient's risk for long-term health issues.[41–43] Robust testing is especially important; some cell lines used in expression systems and PCLs cannot yet be supplied at cGMP quality levels.[33] Genetic assays have already been developed to quantify residual host cell DNA like HEK 293 or HeLa cells by targeting the E1A gene and an L1 retrotransposon, respectively.[43,44] However, novel genetic elements used as *rep*, *cap*, and helper functions may require custom assay design and, since different AAVs production platforms use various culture media, experiments should ensure non-interference. Another major impurity occurs by recombination of rAAV with *rep*, *cap*, and helper elements through fusions between ITRs or when ITRs and promoter regions create replication-competent AAV (rcAAV).[34] This becomes a risk for uncontrolled replication after administration and is also critical to identify and quantify. Using non-related viral components such as HAVE could avoid the production of rcAAV.[45] Genetic tests can also detect residual elements like plasmids that are encapsidated or not fully removed from the matrix of finished products. Analytical analysis to ensure safety and efficacy is critical for the array of other viral and non-viral gene therapy vectors as well. Although this review focuses on assays for AAV vectors, all viral vectors have biological concerns that should be considered during genetic assay development and validation.

Applying PCR-based technologies to AAV therapies

Currently, there is no standardized method for rAAV quantification and depending on the research focus, target virus feature, research interest, and monetary considerations, various analytical methods can be used. PCR assays are generally the most utilized quantitative method for rAAV titration. PCR-based assays are highly standardized with the ability to make direct comparisons of different preparations in terms of assembly or purification efficiency, as well as experimental or therapeutic dosages. Real-time qPCR is based on a fluorescence signal that is emitted by DNA-intercalating dyes or fluorescently labeled target-specific probes that enables the detection and measurement of products generated in each cycle of the amplification process. The success and usage of qPCR in the quantification of rAAVs has been due to factors such as high sensitivity and specificity, low intra- and inter-laboratory variability, wide range of quantification, and simplicity of its performance. Nonetheless, various challenges remain that hinder the robustness and precision in AAV titrations.

The major limitation associated with qPCR in AAV quantification is DNA amplification efficiency, which can be significantly impaired by different factors, such as poor design of primer pairs, presence of inhibitors, or secondary structure in the template, as noted for rcAAV vectors.[46] Additionally, qPCR requires a valid DNA standard curve,[46–48] and accurate calibration of the standard is critical for optimum performance and quantification. For a qPCR method, the number of cycles before fluorescence intensity crosses the threshold is proportional to the amount of target template added to the reaction. As such, there needs to be a calibration curve with dilutions of samples bearing the same target region to calculate the exact copy number of viral genomes. Often the initial target copy number of standards needs to be established using different analytical methods beforehand. Standards used in qPCR assays for AAV titrations are typically either plasmid DNA, linear DNA fragments, purified viral genomes, or actual viruses. For a standard to be considered suitable for the assay, it needs to have certain characteristics such as resembling the secondary structure and stability to obtain precise and reproducible measurements. Different types of amplification errors may result in overestimation or underestimation of rAAV samples.[49] In addition, encapsidated AAV genomes are less accessible than purified plasmid DNA. Finally, the higher order structures in AAV genomes such as

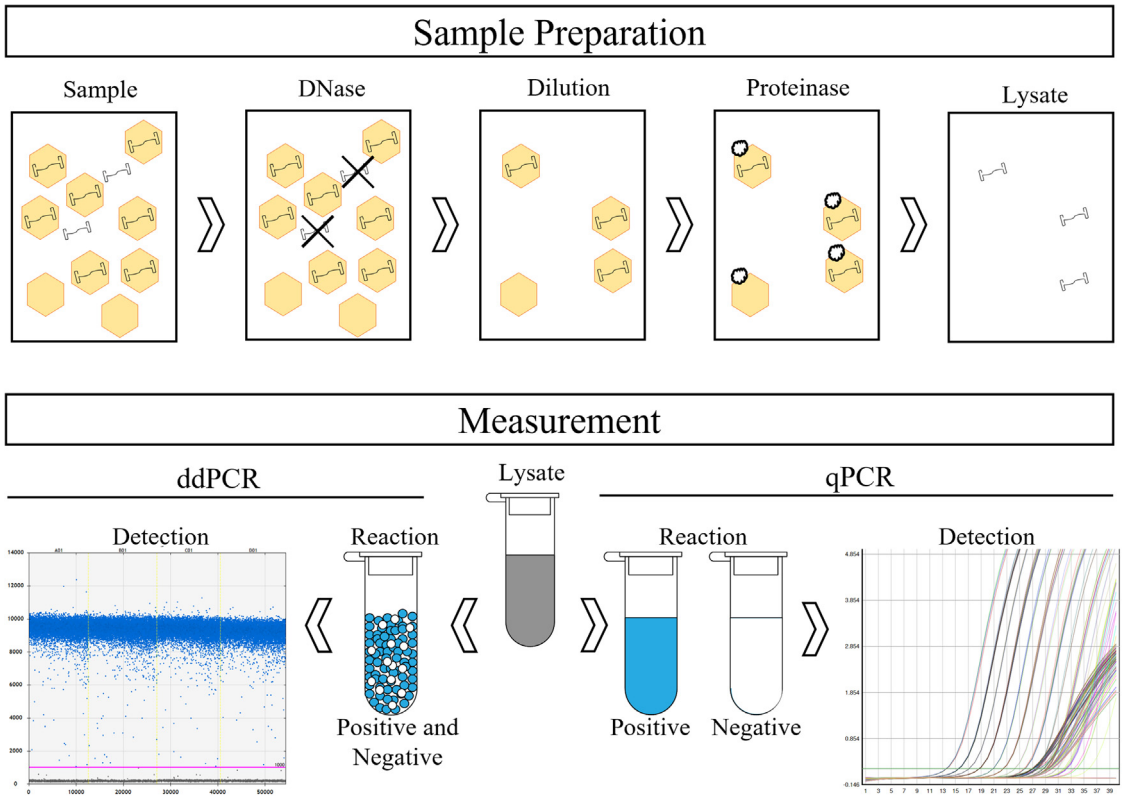

Figure 2. Example of AAV gene therapy sample preparation before PCR analysis

A production sample of rAAV would include capsids with ssDNA, empty capsids, and various residual DNA in the drug matrix. DNase treatment removes DNA in the matrix, ensuring only encapsidated DNA is left. Diluting the sample ensures the sample will be within a method's limit of quantification. Proteinase treatment lyses the capsids, resulting in liberated DNA lysate. The lysate sample is then purified and quantified using qPCR or directly quantified using ddPCR. During qPCR, the fluorophores on DNA oligos excite proportionally to the amount of DNA during thermal cycling. The sample is quantified based on when the signal crosses a base threshold. In ddPCR, the target DNA is diluted into droplets along with the PCR reagents, which results in some droplets being positive or negative. The number of positive versus negative droplets is quantified on a droplet detector.

ITRs on both ends instead of the supercoiled nature of plasmids may affect qPCR efficiency.[50–52] More accurate reference standards for qPCR assays are available, including the use of denatured AAV as a qPCR standard and the use of live AAV as a standard.[52,53]

In recent years, digital PCR has emerged as a technique for the absolute quantification of AAV with the most common example being ddPCR. The commercialization of ddPCR technology became available in 2011 and is similar to qPCR in that both technologies utilize a Taq polymerase in a standard PCR reaction to amplify a target DNA fragment from a complex sample using specific primers and probes.[54,55] However, unlike qPCR, ddPCR has two distinct features. First, the PCR reaction is partitioned into thousands of individual reaction droplets prior to amplification; second, there is the acquisition of data at reaction endpoint instead of in real-time. These two unique features of ddPCR allows for the direct and independent quantification of DNA without standard curves, thus providing more precise, accurate, and reproducible data in comparison with qPCR.[56] Endpoint measurement by ddPCR quantifies nucleic acids independent of the reaction efficiency, thereby resulting in a positive-negative call for every droplet and greater amenability to multiplexed detection of target molecules.[57] This allows for ddPCR technology to be used for extremely low-target quantitation, which would be below the limit of detection using qPCR.

Independent of standards, methods must distinguish vectors that are fully packaged with the functional genome as opposed to empty or truncated particles. The encapsulated AAV genome, rather than AAV ssDNA that may be in the drug matrix, is the key component of the vector that mediates the transfer of the transgene and, therefore, the functional effect. Sample preparation procedures including RNase treatment, proteinase digestion, viral genome extraction, and restriction digestion of ITR coupled with informed primer/probe design have been developed to accurately assess the encapsidated rAAV genomes (Figure 2).[52,58,59] Testing for process impurities depends on the manufacturing process. As mentioned earlier, the various sources of *rep*, *cap*, and helper elements, as well as host cells, represent residual elements that must be quantified. PCR-based assays adhere to the same general parameters no matter the genetic target, yet careful primer and probe design is essential to prevent false

positives and false negatives with complex samples. Additionally, the various culture and buffer ingredients should be considered to avoid potential PCR inhibitors.

Regulatory guidance for developing cGMP methods

Data analysis also presents another layer of complexity by the different sources of DNA from which the samples and standard curves are derived.[60] The publication of "The Minimum Information for the Publication of Quantitative Real-Time PCR Experiments (MIQE)" guidelines and related articles have defined a rigorous methodology for designing, performing, and reporting qPCR and ddPCR experiments that ensures the publication of reproducible and high-quality data.[61–63] There have been numerous retractions of multiple publications that ignored the MIQE-guided protocols over the past several years,[64] which demonstrates the importance of establishing strong method guidelines. Data analysis during cGMP testing must comply with existing requirements for medical data security such as the Code of Federal Regulations, Title 21, Part 11, which was issued in August 1997.[65] Commonly referred to as 21 CFR Part 11, or just Part 11, the regulation sets criteria for electronic record keeping to ensure they are "trustworthy, reliable, and generally equivalent to paper records and handwritten signatures executed on paper."[65] Satisfying Part 11 is a balance between data generation and control of systems, which scales in difficulty depending on the data size and analysis complexity.

While there are many factors for Part 11 compliance that require careful consideration, several criteria are worth highlighting in this review. First, instructions for system use and compliance should be clearly defined by standard operating procedures and trained using a robust learning program to ensure users are knowledgeable and accountable prior to performing a test. Next, access to any system that generates, manipulates, or analyzes data should be controlled to limit access only to users who have undergone appropriate training. Access should require a unique sign-on by the user so that each step in a test is traceable through audit trails and finalized by unique electronic signatures with time stamps. Finally, raw data generated by all systems should automatically be locked down so that an analyst cannot manipulate or delete the records. This usually requires an information technology (IT) strategy of automatically sweeping data to a secure repository. Some PCR systems used in gene therapy testing are initially designed to allow Part 11 compliance or offer specific Part 11 modules that can be implemented in normal operating framework. Other systems that use custom solutions require additional validation before they can be used for cGMP-level testing. As new techniques and technologies are developed to better evaluate therapies, strategies to remain Part 11 compliant must also be implemented. All of these considerations are essential aspects to building robust and reliable methods for cGMP AAV product quantitation.[57]

BUILDING ROBUST MOLECULAR ASSAYS TO QUANTIFY AAV VECTORS AND RESIDUAL ELEMENTS

Recommendations for developing viral genome titer assays

Resources for general analytical test development in the United States are provided by the United States Pharmacopeia (USP), a non-profit, non-government agency that provides monographs, methods, and standards for foods and medicines such as gene therapies.[66] Suggested tests for identity, purity, and potency are explicitly outlined often with specific acceptance criteria in final products. Nucleic acid-based methods like qPCR are suggested for residual host cell DNA, viral identification, labeled dose, and other quality metrics. Regardless of the PCR type, general techniques used in nucleic acid-based techniques, such as extraction, detection, and sequencing, are further outlined in additional USP chapters.[67] Example genetic tests used throughout gene therapy manufacturing and their categories are shown in Table 1. Vector genome titer would be considered quantitation of major components, which would require category I parameters.

Analytical tests must be developed and validated based on best practices depending on the nature of the test. Quantitative assays using PCR will require the method to demonstrate accuracy, precision, specificity, linearity, and range, with a defined upper and lower limit of detection for impurity testing. Method development typically assesses each of these metrics to account for any biological discrepancies that would affect test results, such as those described in the previous sections. USP <1225> is harmonized with The International Council for Harmonization of Technical Requirements for Pharmaceuticals for Humans (ICH) guidance to provide a framework for qualifying or validating methods based on test category and product maturity.[67] In general, products in early phase development only require an assay to be qualified, or shown to fulfill specific conditions.[68,69] Assays for late-phase pharmaceuticals require full validation, meaning the method is shown to be appropriate for use of a specific purpose through sound evidence.[68,69] Table 2 shows an example validation schedule to satisfy the different parameters required for vector genome titer and includes suggested numbers of sample replicates and dilutions, when to use different analysts and instruments, and which parameters can be tested in parallel. The Global Clinical Research Organization Council (GCC) recently published a meta-analysis of qPCR and ddPCR guidance documents that provides recommendation for validating methods.[16] Many of the recommendations are important to discuss in consideration with gene therapy testing.

Accuracy testing is typically performed by spiking plasmids with the transgene or GOI into genomic DNA such as sheared salmon sperm to reduce non-specific reactions. In general, accuracy instances use at least three dilution levels in triplicate. The percent coefficient of variation (%CV) depends on the dilution level, with lower dilutions allowing for higher %CVs. Specificity is contingent on the quality of primer/probe design and optimization, as well as the complexity of the sample. Good practices for primer design are presented in detail by literature and commercial suppliers.[70] Importantly, oligo sequences used for gene therapy methods should be compared *in silico* against any possible genetic elements like *rep*, *cap*, helper, and cell line sequences, as well as other drug product (DP) in the laboratory, to avoid cross-amplification. Negative controls are crucial to ensure run specificity, with non-amplification being less than the limit of

Table 1. USP assay categories and examples for gene therapy testing

		Category II			
	Category I	Quantitative	Limit test	Category III	Category IV
Description					
Category target	quantitation of major components	determination of impurities		determination of performance characteristics	identification tests
Example cGMP test for gene therapy	rAAV genome titer	residual HEK 293 by E1A	residual rcAAV genomes	nuclease activity	rAAV transgene identity
Testing required					
Accuracy	yes	yes	[a]	yes	no
Repeatability	yes	yes	no	yes	no
Intermediate precision	yes	yes	no	yes	no
Specificity	yes	yes	yes	yes	yes
Detection limit	no	no	yes	[a]	no
Quantitation limit	no	yes	no	[a]	no
Linearity	yes	yes	no	yes	no
Range	yes	yes	[a]	yes	no
Robustness	yes	yes	optional	yes	no
System suitability	yes	yes	yes	yes	yes
Standard comparison	optional	optional	no	optional	no
Solution stability	optional	optional	optional	optional	optional

[a]May be required.

detection. Linearity is also a key quality metric in all qPCR methods, since quantification relies on standard curves. The GCC recommends an efficiency range of 90%–100% and a coefficient of determination criterion ($R^2 \geq 0.98$). In sum, all validation parameters should be performed according to current regulatory guidelines while considering the therapy type, sample matrix, and the assay technology to be used.

Although using plasmids to represent GOI for quantification is useful for developing methods, the final validations require full processing of gene therapy DP like encapsidated rAAVs. The general workflow for quantifying encapsidated viral vector genomes in rAAVs by PCR methods includes four discrete steps (Figure 2).

1. Removal of non-encapsidated DNA from viral sample by DNase treatment.
2. Serial dilution of the digested samples.
3. Lysis of the capsid by either heat induction or proteinase K digestion.
4. Quantification of viral vector genomes by qPCR or ddPCR, using a primer/probe set specific to the transgene contained in rAAV sample.

The sample is treated first with DNase I to remove non-encapsidated DNA. Next, depending on the expected concentration of the sample, a pre-dilution step ensures that the sample is within the method's detection range. ddPCR requires a certain number of negative droplets for the Poisson statistics to accurately calculate starting concentration, so it is critical to avoid overloading the reaction.[71] After the lysis, the DNase must be completely inactivated prior to PCR analysis to ensure there is no degradation of newly synthesized DNA. Commonly used methods for removal or inactivation of DNase after digestion include are heat inactivation and proteinase K treatment. Heat inactivation is one of the most common methods of DNase inactivation. Although this method seems to be straightforward, the divalent cations in the DNase digestion buffer can cause (chemically induced) strand scission of RNA when heated. Extensive method development is, thus, required to account for differences in quantification of vector genomes caused by lysis either by heat denaturing or proteinase K lysis. Finally, the lysate is tested with specific primers and analyzed using qPCR or ddPCR. The presence of contaminating genomic DNA in rAAV preparations is a frequent cause of false positives in PCR-based assays aimed at gene expression analysis. PCR primer design can control for genomic DNA contamination by targeting the intron-exon boundaries to amplify a product from contaminating DNA that includes the intron, making it much larger than the expected product size and specific to the target transgene. However, pseudogenes can produce an amplified product of the same size, so it is essential to conduct *in silico* analysis with homologous genes to ensure specificity. Specificity can be further assessed using methods like single droplet population detection using ddPCR or amplicon sequencing using next-generation sequencing (NGS).[72,73] The resulting measurement is the encapsidated, viral genome titer of the target rAAV API.

Table 2. Example validation plan for a qPCR or ddPCR assay of vector genome titer

Requirements	Test 1	Test 2	Test 3	Test 4	Test 5
Analyst	A	A	A	A	B
qPCR/ddPCR instrument	A	A	A	A	B
Day	A	B	C	D	E
No. of replicates of HC 3	1	1	1	–	–
No. of replicates of HC 2	1	1	1	–	–
No. of replicates of HC 1	1	1	1	–	–
No. of replicates of nominal concentration	1	1	1	6	6
No. of replicates of LC 1	1	1	1	–	–
No. of replicates of LC 2	1	1	1	–	–
No. of replicates of LC 3	1	1	1	–	–
System suitability	yes	yes	yes	yes	yes
Specific - non-interference	no	no	yes	no	no
Linearity	yes	yes	yes	no	no
Accuracy	yes	yes	yes	no	no
Repeatability	no	no	no	yes	no
Intermediate precision	no	no	no	yes	yes
Range	yes	yes	yes	yes	no
Quantitation limit	yes	yes	yes	yes	no

HC, high concentration; LC, low concentration.

Residual assay development

The removal of residual host cell proteins and residual plasmids from drug substances is critical in manufacturing high-quality DPs. Extremely sensitive and accurate methods for detection and quantification are needed to accurately quantify these at low levels of detection. In developing methods to quantify residual DNAs, the matrix effect must be addressed; there may be components of the sample that interfere with the assay results. Dilution of the sample may be necessary to overcome the interference, so long as the specified DNA content of the sample remains within the useful range of the analytical procedure. Existing residual DNA detection methods, based on qPCR, typically rely on the extraction of the residual DNA from samples prior to residual DNA quantification to avoid interference of protein and other components. Typically, most extraction approaches use solid-phase DNA extraction procedures based on magnetic particle or liquid phase based on sodium iodide extraction.

Methods to quantify residual elements would fall under category II tests for impurities and require different validation parameters (Table 1). Importantly, detection and quantitation limits must be assessed depending on whether the impurities method is quantitative or a limit test. The pre-processing and DNA purification effects on detection and quantification limits can be assessed by control spiking experiments. All method categories also require proper system suitability parameters using various controls to distinguish between system failures or sample irregularities. Methods to quantify residual elements must be validated for the parameters of category II assays. The development and validation process should be repeated for use in each new API, even if the target residual is the same. For example, a validated assay to target residual HEK 293 DNA in one gene therapy requires re-validation if the method is to be used in a new therapy. Altogether, regulations for developing robust genetic tests are available to allow researchers to apply tests for new therapies or develop and validate new tests to improve the industry.

OPPORTUNITIES FOR IMPROVING ANALYTICAL TESTING

Reducing development efforts

The rigorous requirements of analytical method development demonstrated throughout this review are crucial to ensuring clinical effects in manufactured batches of therapies. There are many opportunities for improving nucleic acid-based analytical testing to increase speed and decrease costs. The first, and most obvious, is to do less overall work through the development of platform methods. Instead of designing each method *à la carte*, a platform method can be used to encompass a defined category of therapy. The example viral genome titer above describes the quantitation of the transgene in an AAV gene therapy DP. If a different transgene is programmed into the same vector, and the production parameters are exactly equivalent, a new method for vector quantification and residual analysis would typically be developed and validated *ab initio*. This is important to account for unforeseen difficulties caused by the physical DNA sequence. For instance, without evaluation, regions that are G+C rich or that have uncharacterized repeats could interfere with the PCR reaction, either in decreasing the efficiency or obfuscating the overall reaction. If a qPCR or ddPCR method is validated to detect residual concentrations of a specific helper plasmid, and the same plasmid is used to produce an AAV drug substances (DSs) with different transgenes, theoretically, the same method could be used in the different DSs. The difference in transgene sequence would not likely affect the PCR efficiency and quantitative power of the assay of the helper plasmid. Importantly, the other reagents in the manufacturing and storage buffer would also have to be assessed for equivalency to ensure specificity against background interference. Applying this platform approach could decrease the overall numbers of methods and time required for cGMP-compliant manufacturing. Moreover, as bioinformatics improve, rapid *in silico* evaluation can be used to evaluate APIs for potential sequence issues and decrease the laboratory work needed to optimize methods.

The often-unused capability to multiplex assays in PCR is another opportunity to decrease workload. For example, the potency of an rAAV therapy is not solely dependent on the AAV titer, but also relies on the vector genome integrity. This implies that there cannot be a full determination of the potency of an AAV therapy without accurate assessment of both the titer of the viral genome and the concentration of the infectious genomes.[74] A manufactured batch may contain degraded products, contaminant DNA, or truncated vector genomes, and it can be challenging to differentiate between these entities and full, functional vector genomes. This can lead to the production of

therapies that may not contain enough infectious genomes thus decreasing the efficacy and effectiveness of the therapy. To overcome this issue, there has been suggestion of the use of two-dimensional (2D) ddPCR of fluorescein- and hexachloro-6-carboxy-fluorescien-labeled probes targeting different positions of the same rAAV genome.[75] The 2D ddPCR was also shown to be effective in the evaluation of plasmid DNA impurities utilized for vector production. The assessment of different targets using the multiplexing approach would not replace the potency, expression, or infectious titer assays typically used in rAAV quantification, but rather supplement safety and quality profiles of the rAAV therapies. In addition, multiplexing can identify multiple rAAV variants within a single sample when using high-throughput methods such as NGS.[76] Although promising, these multiplexing methods have not been widely utilized in the industry settings and further considerations and feasibility studies need to be conducted in respect of designing an analytical method that meets the requirements of the ICH guidelines.

Improving genetic assays for use with non-genetic elements

Non-viral-based DSs are also important options for delivering gene therapies.[5,77] Vectors based on peptides, lipids, polymers, and inorganic materials expand the repertoire of tropisms for gene therapy and offer enhanced safety and efficacy profiles compared with viruses. A compelling example is mRNA vaccines delivered by lipid nanoparticles that were critical in the response to the severe acute respiratory syndrome coronavirus 2 pandemic.[78] However, these vectors present novel challenges when using PCR and other genetic assays to characterize associated APIs. For example, polyethyleneimine, a biopolymer used as a gene therapy vector,[79] was shown to affect the electrostatic relationship of PCR components and even enhance non-specific PCR reactions.[80] Many non-viral vectors require de-formulation processes to liberate nucleic acid APIs before genetic assays can proceed and some detergents used for de-formulation may inhibit PCR reactions.[81] These artificial enhancements or inhibition to PCR results could drastically affect final measurements since raw data are usually multiplied by several dilution factors. Therefore, APIs should be purified after de-formulation to prevent affecting downstream processes like reverse transcription of mRNA to cDNA and PCR fidelity. Conducting stringent feasibility and validation studies is required to identify and solve issues caused by non-viral vector components.

Non-genetic vectors and residual chemical elements are often characterized by traditional non-genetic methods such as high-performance liquid chromatography, capillary gel electrophoresis (CGE), and mass spectrometry (MS). Interestingly, these methods can be leveraged to improve the development of genetic assays like qPCR or sequencing. An example of this is with mRNAs, which typically have a 5′ cap, a 5′ UTR, a coding region composed of a unique sequence for the target disease or treatment, a 3′ UTR, and a poly-adenylated (poly(A)) tail.[82] Identifying the end-to-end sequence of the mRNA by a method like DNA sequencing may be difficult since the homopolymer region of the poly(A) tail would likely induce sequence error from polymerase slippage.[83] Alternatively, CGE is able to characterize the length of mRNA poly(A) tails,[84] while methods like liquid chromatography-MS can analyze the 5′ cap.[85] These methods would enable researchers to focus on streamlining the DNA sequencing method to only identify the UTRs and coding regions, a much easier endeavor. Thus, it is imperative for researchers to be aware of the array of genetic and non-genetic analytical methods to form strategies for the best ways to assess the various constituents of vectors and APIs.

NGS for gene therapy

NGS is another exciting opportunity for the improvement for cGMP analytics. NGS refers to large-scale, rapid, and efficient DNA (or RNA) sequencing technology. Integration of NGS into clinical drug development has the potential to accelerate clinical trial conduct and improve clinical patient care landscape by identifying rare diseases in patients sooner using less invasive treatment options. Certain strategies of NGS can determine the sequence of unknown samples, as opposed to PCR which requires the target sequence to be known for primer and probe design. So-called shotgun sequencing enables searching for events that caused by to random recombinations, such as the generation of rcAAVs.[86] The general method to check for rcAAV involves subsequent culturing in host cells, followed by assessment of increasing titers using a method like ddPCR.[87] This process is time intensive, often taking several days to weeks to generate reliable data, whereas using NGS on the entire batch would allow rapid identification of cross-over events. Moreover, rcAAV genome variability can be captured in individual reads using long read sequencing.[34] This avoids the targeting requirement of PCR and provides a broad assessment of recombinant targets.

In addition to detecting unknown targets, using NGS for gene therapy analytics is an opportunity for platform methods and multiplexing as well. Because bulk DNA in a sample is sequenced using the same library preparation techniques, little development is needed for similar therapies once a process is in place. In fact, robust platform processes have already been developed to identify genetic elements like AAV genomes and associated plasmids.[34,88] Multiplexing is regularly used in NGS to take advantage of the massive sequencing output generated by new technologies. Sample DNA libraries can be prepared using various methods for plasmid, viral, or eukaryotic host cell procedures; barcoded; and then run altogether in a single sequencing run. The sequencing depth provides such a large number of reads that target quantitation can be performed. This is commonly used to measure other biological investigations like cancer allele measurement and microbiome analysis.[89–91] Microbiome studies assessing 16S rRNA gene sequences of microbial communities are usually presented in relative amounts, although absolute quantitation is possible using controls with known amounts of bacterial genomes.[92] Assay development steps outlined by USP and ICH for quantitative assays will need to be consulted if NGS is to be used for quantitation assays for cGMP gene therapy manufacturing.

There are other important considerations to ensure quantitative NGS assays are compliant for pharmacogenomic use, which is being evaluated by the MicroArray/Sequencing Quality Control consortium at the FDA.[93] These efforts will help to identify quality metrics and

criteria, data analysis protocols, and other effects of NGS results.[94] Per 21 CFR part 11, the massive data output from each sequencing run must be appropriately saved and secured. Robust IT support is required to create an automated process to sweep data directly from the sequencer to a data repository, whether in a local mainframe or in a cloud-based solution. Raw data must be secured against data loss or modification, which could be an issue for experiments that can generate terabytes of raw data.[95] Another challenge is establishing requirements for high-quality sequencing, since there is not well-published regulatory guidance. Although tools to assess sequencing runs are available to provide key metrics such as Phred scores, regulatory guidance is needed for analytical testing during gene therapy manufacturing using NGS. The last consideration for applying NGS is the process of analyzing sequencing data, which typically requires complex, custom bioinformatics programs for each step of data analysis. As discussed above, a cGMP test method must be secured and standardized to remove as much variability from the procedure, so NGS experiments for gene therapies will require robust yet regimented *in silico* analysis procedures using programs that have single-user logins and audit support for 21 CFR part 11 compliance. Despite these challenges, NGS is an important technology that can provide more rapid, deeper insights into gene therapy products than current genetic methods.

Using artificial intelligence for GMP method development

Another exciting field with potential opportunities to improve medicine is the use of artificial intelligence (AI).[96] Recent tools like ChatGPT and BioGPT are a few mainstream examples of AI infused technology.[97,98] In gene therapy, AI could conduct predictions on the complex interactions in drug performance and immune responses to vastly improve drug discovery.[99,100] AI could also assist with complex topics like analyzing many of the genomic issues that have been discussed in this review,[101,102] or more relatively simple topics like electronic documentation best practices. Indeed, AI has already been used to assist with medical reporting for radiology and cancer.[103,104] This holds promise in analytical laboratories, where methods and protocols can sometimes include dozens to hundreds of pages, which are multiplied by each new therapy. AI can take input data to predict text and generate functional methods very rapidly.[105] AI tools for translating languages can also be leveraged for methods shared globally to increase reproducibility between laboratories.[106] Additionally, data generated during cGMP testing of new batches and stability must be closely reviewed and any discrepancies require quality investigations. AI tools can proofread and review analytical data to detect errors and conduct investigations.[107] Using AI simply for improving document generation and review during analytical testing of gene therapies could quickly impact the development time and quality; however, there are several important safety and technical aspects that remain to be solved before widespread use.

The technical and ethical challenges of implementing AI in the medical field have been extensively reviewed,[108,109] which include considerations for application during cGMP method development. AI systems are forward thinking and only as good as the input data used to train it. Massive amounts of curated data must be available to train any AI system before it is considered reliable. The cost of generating the datasets, training programs, reviewing output, and upkeep is a high barrier to entry. AI systems will still likely require auditing by specialized reviewers to ensure accountability and prevent implicit biases that could arise from uneven data used for training.[109] Throughout the process data hygiene would be paramount to ensure patient and proprietary information remains secure. Notwithstanding these inherent challenges, AI is an opportunity to revolutionize cGMP analysis of new gene therapies in the future.

CONCLUSIONS

Experts agree that manufacturing is one of the primary bottlenecks toward widespread application of gene therapies.[110] Investigations into the biological and molecular attributes of delivery vectors like AAVs have allowed manufacturers to increase yields and reduce production time. Manufacturers must also continue to ensure batches meet quality requirements, so therapies are safe and efficient for patients. Analytical methods for new therapies must be quickly validated and implemented while accounting for the breadth of technologies used in manufacturing and remaining cost-effective. In response, genetic assays have progressed rapidly to quantify target therapies and residual effectors efficiently and resolutely, but there are still many opportunities to streamline the development process. Indeed, NGS is already becoming a popular and necessary tool to quickly detect safety issues like replication competent AAV genomes. In summary, researchers must ensure that the computational and regulatory underpinnings of analytical methods keep up with the rate of wet laboratory method improvements to ensure the reliable production of life changing gene therapies.

ACKNOWLEDGMENTS

The authors thank Tomas Gverzdys, Katherine Hanson, Sangita Suresh, Marc Wolman, and Caleb York at PPD for their critical review of the manuscript. We also want to thank Scott Amaya for the graphical design of Figure 1. This work was funded by PPD, the Clinical Research Group of ThermoFisher Scientific.

AUTHOR CONTRIBUTIONS

E. B. and W.M. led manuscript drafting. All authors contributed to and reviewed the final manuscript.

DECLARATION OF INTERESTS

At the time of writing, all authors were employees of the Gene & Cell Therapy unit, PPD GMP Laboratories, part of ThermoFisher Scientific, which develops methods and performs analytical testing for pharmaceuticals including gene and cell therapies.

REFERENCES

1. Pupo, A., Fernández, A., Low, S.H., François, A., Suárez-Amarán, L., and Samulski, R.J. (2022). AAV vectors: The Rubik's cube of human gene therapy. Mol. Ther. *30*, 3515–3541.
2. Therapy, A.S.G.C.T. (2023). Gene, Cell, & RNA Therapy Landscape: Q1 2023 Quarterly Data Report.

3. FDA, U. (2020). Long term follow-up after administration of human gene therapy products. Guidance for Industry.
4. Maeder, M.L., and Gersbach, C.A. (2016). Genome-editing technologies for gene and cell therapy. Mol. Ther. *24*, 430–446.
5. Zu, H., and Gao, D. (2021). Non-viral vectors in gene therapy: Recent development, challenges, and prospects. AAPS J. *23*, 78.
6. Bulcha, J.T., Wang, Y., Ma, H., Tai, P.W.L., and Gao, G. (2021). Viral vector platforms within the gene therapy landscape. Signal Transduct. Target. Ther. *6*, 53.
7. FDA, U. (1978). Title 21 of the US Code of Federal Regulations: 21 CFR 58 Good Laboratory Practice for Nonclinical Laboratory Studies.
8. FDA, U. (1980). Title 21 of the US Code of Federal Regulations: 21 CFR 50 Protection of Human Subjects.
9. FDA, U. (1978). Title 21 of the US Code of Federal Regulations: 21 CFR 211 current good manufacturing practice for finished pharmaceuticals.
10. Research, C.f.B.E.a. (2022). Human Gene Therapy Products Incorporating Human Genome Editing.
11. Flotte, T.R., Trapnell, B.C., Humphries, M., Carey, B., Calcedo, R., Rouhani, F., Campbell-Thompson, M., Yachnis, A.T., Sandhaus, R.A., McElvaney, N.G., et al. (2011). Phase 2 clinical trial of a recombinant adeno-associated viral vector expressing α1-antitrypsin: interim results. Hum. Gene Ther. *22*, 1239–1247.
12. McIntosh, J.H., Cochrane, M., Cobbold, S., Waldmann, H., Nathwani, S.A., Davidoff, A.M., and Nathwani, A.C. (2012). Successful attenuation of humoral immunity to viral capsid and transgenic protein following AAV-mediated gene transfer with a non-depleting CD4 antibody and cyclosporine. Gene Ther. *19*, 78–85.
13. Venditti, C.P. (2021). Safety questions for AAV gene therapy. Nat. Biotechnol. *39*, 24–26.
14. Perez, B.A., Shutterly, A., Chan, Y.K., Byrne, B.J., and Corti, M. (2020). Management of neuroinflammatory responses to AAV-mediated gene therapies for neurodegenerative diseases. Brain Sci. *10*, 119.
15. Flotte, T.R. (2020). Revisiting the "New" Inflammatory Toxicities of Adeno-Associated Virus Vectors (Mary Ann Liebert, Inc).
16. 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.
17. 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.
18. Lundstrom, K. (2018). Viral vectors in gene therapy. Diseases *6*, 42.
19. Robbins, P.D., Ghivizzani, S.C., and Ghivizzani, S.C. (1998). Viral vectors for gene therapy. Pharmacol. Ther. *80*, 35–47.
20. Osten, P., Grinevich, V., and Cetin, A. (2007). Viral vectors: a wide range of choices and high levels of service. In Conditional Mutagenesis: An Approach to Disease Models, R. Feil and D. Metzger, eds. (Springer), pp. 177–202.
21. Ghosh, S., Brown, A.M., Jenkins, C., and Campbell, K. (2020). Viral vector systems for gene therapy: a comprehensive literature review of progress and biosafety challenges. Appl. Biosaf. *25*, 7–18.
22. Shmidt, A.A., and Egorova, T.V. (2021). PCR-based analytical methods for quantification and quality control of recombinant adeno-associated viral vector preparations. Pharmaceuticals *15*, 23.
23. Pellenz, S., Phelps, M., Tang, W., Hovde, B.T., Sinit, R.B., Fu, W., Li, H., Chen, E., and Monnat, R.J., Jr. (2019). New human chromosomal sites with "safe harbor" potential for targeted transgene insertion. Hum. Gene Ther. *30*, 814–828.
24. Warnock, J.N., Daigre, C., and Al-Rubeai, M. (2011). Introduction to viral vectors. In Viral Vectors for Gene Therapy: Methods and Protocols, O.-W. Merten and M. Al-Rubeai, eds. (Springer), pp. 1–25.
25. Nóbrega, C., Mendonça, L., and Matos, C.A. (2020). A Handbook of Gene and Cell Therapy (Springer).
26. Balakrishnan, B., and Jayandharan, G.R. (2014). Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. Curr. Gene Ther. *14*, 86–100.
27. Meier, A.F., Fraefel, C., and Seyffert, M. (2020). The interplay between adeno-associated virus and its helper viruses. Viruses *12*, 662.
28. Samulski, R.J., and Muzyczka, N. (2014). AAV-mediated gene therapy for research and therapeutic purposes. Annu. Rev. Virol. *1*, 427–451.
29. Pereira, D.J., McCarty, D.M., and Muzyczka, N. (1997). The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. J. Virol. *71*, 1079–1088.
30. Linden, R.M., and Berns, K.I. (2000). Molecular biology of adeno-associated viruses. Contrib. Microbiol. *4*, 68–84.
31. McCarty, D.M. (2008). Self-complementary AAV vectors; advances and applications. Mol. Ther. *16*, 1648–1656.
32. Naso, M.F., Tomkowicz, B., Perry, W.L., III, and Strohl, W.R. (2017). Adeno-associated virus (AAV) as a vector for gene therapy. BioDrugs *31*, 317–334.
33. Dobrowsky, T., Gianni, D., Pieracci, J., and Suh, J. (2021). AAV manufacturing for clinical use: insights on current challenges from the upstream process perspective. Curr. Opin. Biomed. Eng. *20*, 100353.
34. Radukic, M.T., Brandt, D., Haak, M., Müller, K.M., and Kalinowski, J. (2020). Nanopore sequencing of native adeno-associated virus (AAV) single-stranded DNA using a transposase-based rapid protocol. NAR Genom. Bioinform. *2*, lqaa074.
35. van Oers, M.M. (2011). Opportunities and challenges for the baculovirus expression system. J. Invertebr. Pathol. *107*, S3–S15.
36. Jorio, H., Tran, R., and Kamen, A. (2006). Stability of serum-free and purified baculovirus stocks under various storage conditions. Biotechnol. Prog. *22*, 319–325.
37. Selvaraj, N., Wang, C.-K., Bowser, B., Broadt, T., Shaban, S., Burns, J., Saptharishi, N., Pechan, P., Golebiowski, D., Alimardanov, A., et al. (2021). Detailed protocol for the novel and scalable viral vector upstream process for AAV gene therapy manufacturing. Hum. Gene Ther. *32*, 850–861.
38. Aponte-Ubillus, J.J., Barajas, D., Peltier, J., Bardliving, C., Shamlou, P., and Gold, D. (2018). Molecular design for recombinant adeno-associated virus (rAAV) vector production. Appl. Microbiol. Biotechnol. *102*, 1045–1054.
39. Martin, J., Frederick, A., Luo, Y., Jackson, R., Joubert, M., Sol, B., Poulin, F., Pastor, E., Armentano, D., Wadsworth, S., and Vincent, K. (2013). Generation and characterization of adeno-associated virus producer cell lines for research and preclinical vector production. Hum. Gene Ther. Methods *24*, 253–269.
40. Su, W., Patrício, M.I., Duffy, M.R., Krakowiak, J.M., Seymour, L.W., and Cawood, R. (2022). Self-attenuating adenovirus enables production of recombinant adeno-associated virus for high manufacturing yield without contamination. Nat. Commun. *13*, 1182.
41. Xu, J., DeVries, S.H., and Zhu, Y. (2020). Quantification of adeno-associated virus with safe nucleic acid dyes. Hum. Gene Ther. *31*, 1086–1099.
42. Sheng, L., Cai, F., Zhu, Y., Pal, A., Athanasiou, M., Orrison, B., Blair, D.G., Hughes, S.H., Coffin, J.M., Lewis, A.M., and Peden, K. (2008). Oncogenicity of DNA in vivo: Tumor induction with expression plasmids for activated H-ras and c-myc. Biologicals *36*, 184–197.
43. Zheng, W., Jiang, L., Lei, Q., Yang, J., Gao, X., Wang, W., Zhang, Y., Kong, T., Chen, Q., and Li, G. (2019). Development and validation of quantitative real-time PCR for the detection of residual CHO host cell DNA and optimization of sample pretreatment method in biopharmaceutical products. Biol. Proced. Online *21*, 17.
44. Rahbari, R., Sheahan, T., Modes, V., Collier, P., Macfarlane, C., and Badge, R.M. (2009). A novel L1 retrotransposon marker for HeLa cell line identification. Biotechniques *46*, 277–284.
45. Clément, N., Knop, D.R., and Byrne, B.J. (2009). Large-scale adeno-associated viral vector production using a herpesvirus-based system enables manufacturing for clinical studies. Hum. Gene Ther. *20*, 796–806.
46. Fagone, P., Wright, J.F., Nathwani, A.C., Nienhuis, A.W., Davidoff, A.M., and Gray, J.T. (2012). Systemic errors in quantitative PCR titration of self-complementary AAV vectors and improved alternative methods. Hum. Gene Ther. Methods *23*, 1–7.
47. d'Costa, S., Blouin, V., Broucque, F., Penaud-Budloo, M., François, A., Perez, I.C., Le Bec, C., Moullier, P., Snyder, R.O., and Ayuso, E. (2016). Practical utilization of recombinant AAV vector reference standards: focus on vector genomes titration by free ITR qPCR. Mol. Ther. Methods Clin. Dev. *5*, 16019.

48. 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.

49. Mayginnes, J.P., Reed, S.E., Berg, H.G., Staley, E.M., Pintel, D.J., and Tullis, G.E. (2006). Quantitation of encapsidated recombinant adeno-associated virus DNA in crude cell lysates and tissue culture medium by quantitative, real-time PCR. J. Virol. Methods *137*, 193–204.

50. Martinez-Fernandez de la Camara, C., McClements, M.E., and MacLaren, R.E. (2021). Accurate quantification of AAV vector genomes by quantitative PCR. Genes *12*, 601.

51. Wang, F., Cui, X., Wang, M., Xiao, W., Xu, R., and Xu, R. (2013). A reliable and feasible qPCR strategy for titrating AAV vectors. Med. Sci. Monit. Basic Res. *19*, 187–193.

52. Wang, Y., Menon, N., Shen, S., Feschenko, M., and Bergelson, S. (2020). A qPCR method for AAV genome titer with ddPCR-level of accuracy and precision. Mol. Ther. Methods Clin. Dev. *19*, 341–346.

53. Dobnik, D., Kogovšek, P., Jakomin, T., Košir, N., Tušek Žnidarič, M., Leskovec, M., Kaminsky, S.M., Mostrom, J., Lee, H., and Ravnikar, M. (2019). Accurate quantification and characterization of adeno-associated viral vectors. Front. Microbiol. *10*, 1570.

54. Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero, M.Y., Hiddessen, A.L., Legler, T.C., et al. (2011). High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal. Chem. *83*, 8604–8610.

55. Hindson, C.M., Chevillet, J.R., Briggs, H.A., Gallichotte, E.N., Ruf, I.K., Hindson, B.J., Vessella, R.L., and Tewari, M. (2013). Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat. Methods *10*, 1003–1005.

56. Pacouret, S., Bouzelha, M., Shelke, R., Andres-Mateos, E., Xiao, R., Maurer, A., Mevel, M., Turunen, H., Barungi, T., Penaud-Budloo, M., et al. (2017). AAV-ID: a rapid and robust assay for batch-to-batch consistency evaluation of AAV preparations. Mol. Ther. *25*, 1375–1386.

57. Taylor, S.C., Laperriere, G., and Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. Sci. Rep. *7*, 2409.

58. Ai, J., Ibraheim, R., Tai, P.W.L., and Gao, G. (2017). A scalable and accurate method for quantifying vector genomes of recombinant adeno-associated viruses in crude lysate. Hum. Gene Ther. Methods *28*, 139–147.

59. Werling, N.J., Satkunanathan, S., Thorpe, R., and Zhao, Y. (2015). Systematic comparison and validation of quantitative real-time PCR methods for the quantitation of adeno-associated viral products. Hum. Gene Ther. Methods *26*, 82–92.

60. Carvalhais, V., Delgado-Rastrollo, M., Melo, L.D.R., and Cerca, N. (2013). Controlled RNA contamination and degradation and its impact on qPCR gene expression in *S. epidermidis* biofilms. J. Microbiol. Methods *95*, 195–200.

61. 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: M Inimum I Nformation for Publication of Q Uantitative Real-Time PCR E Xperiments (Oxford University Press).

62. Taylor, S., Wakem, M., Dijkman, G., Alsarraj, M., and Nguyen, M. (2010). A practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines. Methods *50*, S1–S5.

63. Huggett, J.F., Foy, C.A., Benes, V., Emslie, K., Garson, J.A., Haynes, R., Hellemans, J., Kubista, M., Mueller, R.D., Nolan, T., et al. (2013). The Digital MIQE Guidelines: M inimum I nformation for Publication of Q uantitative Digital PCR E xperiments. Clin. Chem. *59*, 892–902.

64. Bustin, S.A., and Nolan, T. (2016). Improving the Reliability of Peer-Reviewed Publications: We Are All in it Together (Elsevier).

65. FDA, U. (1997). Title 21 CFR Part 11 - Electronic Records; Electronic Signatures.

66. USP (2020). <1047> Gene Therapy Products. Pharmacopeial Forum 44. https://doi.org/10.31003/USPNF_M3024_02_01.

67. USP (2017). <1225> Validation of Compendial Procedures. Pharmacopeial Forum 42. https://doi.org/10.31003/USPNF_M99945_04_01.

68. Iso, E. (2005). 9000: 2005. Quality Management Systems-Fundamentals and Vocabulary (ISO 9000: 2005) 1.

69. Salsbury, B. (2012). Defining Qualification, Verification, and Validation. In Ask the Standards Experts American Society for Quality Online Forum.

70. Bustin, S., and Huggett, J. (2017). qPCR primer design revisited. Biomol. Detect. Quantif. *14*, 19–28.

71. Hansen, S.J.Z., Tang, P., Kiefer, A., Galles, K., Wong, C., and Morovic, W. (2019). Droplet digital PCR is an improved alternative method for high-quality enumeration of viable probiotic strains. Front. Microbiol. *10*, 3025.

72. Miotke, L., Lau, B.T., Rumma, R.T., and Ji, H.P. (2014). High sensitivity detection and quantitation of DNA copy number and single nucleotide variants with single color droplet digital PCR. Anal. Chem. *86*, 2618–2624.

73. Saingam, P., Li, B., and Yan, T. (2018). Use of amplicon sequencing to improve sensitivity in PCR-based detection of microbial pathogen in environmental samples. J. Microbiol. Methods *149*, 73–79.

74. Ayuso, E., Blouin, V., Lock, M., Mcgorray, S., Leon, X., Alvira, M.R., Auricchio, A., Bucher, S., Chtarto, A., Clark, K.R., et al. (2014). Manufacturing and characterization of a recombinant adeno-associated virus type 8 reference standard material. Hum. Gene Ther. *25*, 977–987.

75. Furuta-Hanawa, B., Yamaguchi, T., and Uchida, E. (2019). Two-dimensional droplet digital PCR as a tool for titration and integrity evaluation of recombinant adeno-associated viral vectors. Hum. Gene Ther. Methods *30*, 127–136.

76. Weinmann, J., Weis, S., Sippel, J., Tulalamba, W., Remes, A., El Andari, J., Herrmann, A.-K., Pham, Q.H., Borowski, C., Hille, S., et al. (2020). Identification of a myotropic AAV by massively parallel in vivo evaluation of barcoded capsid variants. Nat. Commun. *11*, 5432.

77. Ramamoorth, M., and Narvekar, A. (2015). Non viral vectors in gene therapy-an overview. J. Clin. Diagn. Res. *9*, GE01–GE06.

78. Labouta, H.I., Langer, R., Cullis, P.R., Merkel, O.M., Prausnitz, M.R., Gomaa, Y., Nogueira, S.S., and Kumeria, T. (2022). Role of drug delivery technologies in the success of COVID-19 vaccines: a perspective. Drug Deliv. Transl. Res. *12*, 2581–2588.

79. Ryu, N., Kim, M.-A., Park, D., Lee, B., Kim, Y.-R., Kim, K.-H., Baek, J.-I., Kim, W.J., Lee, K.-Y., and Kim, U.-K. (2018). Effective PEI-mediated delivery of CRISPR-Cas9 complex for targeted gene therapy. Nanomedicine *14*, 2095–2102.

80. Tong, W., Cao, X., Wen, S., Guo, R., Shen, M., Wang, J., and Shi, X. (2012). Enhancing the specificity and efficiency of polymerase chain reaction using polyethyleneimine-based derivatives and hybrid nanocomposites. Int. J. Nanomed. *7*, 1069–1078.

81. Schrader, C., Schielke, A., Ellerbroek, L., and Johne, R. (2012). PCR inhibitors–occurrence, properties and removal. J. Appl. Microbiol. *113*, 1014–1026.

82. Chaudhary, N., Weissman, D., and Whitehead, K.A. (2021). mRNA vaccines for infectious diseases: principles, delivery and clinical translation. Nat. Rev. Drug Discov. *20*, 817–838.

83. Kircher, M., and Kelso, J. (2010). High-throughput DNA sequencing–concepts and limitations. Bioessays *32*, 524–536.

84. Warzak, D.A., Pike, W.A., and Luttgeharm, K.D. (2023). Capillary Electrophoresis Methods for Determining the IVT mRNA Critical Quality Attributes of Size and Purity (SLAS Technology).

85. Galloway, A., Atrih, A., Grzela, R., Darzynkiewicz, E., Ferguson, M.A.J., and Cowling, V.H. (2020). CAP-MAP: cap analysis protocol with minimal analyte processing, a rapid and sensitive approach to analysing mRNA cap structures. Open Biol. *10*, 190306.

86. Allen, J.M., Debelak, D.J., Reynolds, T.C., and Miller, A.D. (1997). Identification and elimination of replication-competent adeno-associated virus (AAV) that can arise by nonhomologous recombination during AAV vector production. J. Virol. *71*, 6816–6822.

87. Dash, S., Sharon, D.M., Mullick, A., and Kamen, A.A. (2022). Only a small fraction of cells produce assembled capsids during transfection-based manufacturing of adeno-associated virus vectors. Biotechnol. Bioeng. *119*, 1685–1690.

88. Namkung, S., Tran, N.T., Manokaran, S., He, R., Su, Q., Xie, J., Gao, G., and Tai, P.W.L. (2022). Direct ITR-to-ITR nanopore sequencing of AAV vector genomes. Hum. Gene Ther. *33*, 1187–1196.

89. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. *37*, 852–857.

90. Serratì, S., De Summa, S., Pilato, B., Petriella, D., Lacalamita, R., Tommasi, S., and Pinto, R. (2016). Next-generation sequencing: advances and applications in cancer diagnosis. OncoTargets Ther. *9*, 7355–7365.

91. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

92. Tkacz, A., Hortala, M., and Poole, P.S. (2018). Absolute quantitation of microbiota abundance in environmental samples. Microbiome *6*, 110–113.

93. SEQC/MAQC-III Consortium (2014). A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat. Biotechnol. *32*, 903–914.

94. FDA, U. (2019). MicroArray/Sequencing Quality Control (MAQC/SEQC).

95. Schmidt, B., and Hildebrandt, A. (2017). Next-generation sequencing: big data meets high performance computing. Drug Discov. Today *22*, 712–717.

96. Briganti, G., and Le Moine, O. (2020). Artificial intelligence in medicine: today and tomorrow. Front. Med. *7*, 27.

97. OpenAI, T. (2022). Chatgpt: Optimizing Language Models for Dialogue (OpenAI).

98. Luo, R., Sun, L., Xia, Y., Qin, T., Zhang, S., Poon, H., and Liu, T.-Y. (2022). BioGPT: generative pre-trained transformer for biomedical text generation and mining. Brief. Bioinform. *23*, bbac409.

99. Hasanzadeh, A., Hamblin, M.R., Kiani, J., Noori, H., Hardie, J.M., Karimi, M., and Shafiee, H. (2022). Could artificial intelligence revolutionize the development of nanovectors for gene therapy and mRNA vaccines? Nano Today *47*, 101665.

100. Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., and Tekade, R.K. (2021). Artificial intelligence in drug discovery and development. Drug Discov. Today *26*, 80–93.

101. Caudai, C., Galizia, A., Geraci, F., Le Pera, L., Morea, V., Salerno, E., Via, A., and Colombo, T. (2021). AI applications in functional genomics. Comput. Struct. Biotechnol. J. *19*, 5762–5790.

102. Dias, R., and Torkamani, A. (2019). Artificial intelligence in clinical and genomic diagnostics. Genome Med. *11*, 70.

103. Jeblick, K., Schachtner, B., Dexl, J., Mittermeier, A., Stüber, A.T., Topalis, J., Weber, T., Wesp, P., Sabel, B., Ricke, J., et al. (2022). ChatGPT Makes Medicine Easy to Swallow: An Exploratory Case Study on Simplified Radiology Reports. Preprint at arXiv. https://doi.org/10.48550/arXiv.2212.14882.

104. Chen, S., Kann, B.H., Foote, M.B., Aerts, H.J., Savova, G.K., Mak, R.H., and Bitterman, D.S. (2023). The utility of ChatGPT for cancer treatment information. Preprint at medRxiv. https://doi.org/10.1101/2023.03.16.23287316.

105. Lee, D.-G., and Seo, Y.-S. (2020). Improving bug report triage performance using artificial intelligence based document generation model. Hum. Cent. Comput. Inf. Sci. *10*, 26.

106. Tsai, S.-C. (2019). Using google translate in EFL drafts: a preliminary investigation. Comput. Assist. Lang. Learn. *32*, 510–526.

107. Zhao, H. (2021). Design of Intelligent Proofreading System Based on Artificial Intelligence. 2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture (AIAM) (IEEE).

108. Rajpurkar, P., Chen, E., Banerjee, O., and Topol, E.J. (2022). AI in health and medicine. Nat. Med. *28*, 31–38.

109. Gerke, S., Minssen, T., and Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In Artificial Intelligence in Healthcare (Elsevier), pp. 295–336.

110. Wang, D., Tai, P.W.L., and Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. Nat. Rev. Drug Discov. *18*, 358–378.