ddPCR-Based Quantification and Fragment Analysis of Residual E1A in rAAV Products: Method Validation and Correlation with Total Host Cell DNA

María García[1], Ainara Arana[1], Alba Judith Mateos[1], Marina Barriocanal[1], Adrián Perez-San Vicente[1], Oihana Aldasoro[1], Ilya Nechaev[1], Achille Francois[1], César Trigueros[1]
[1]Viralgen Vector Core S.L., San Sebastián, Gipuzkoa, 20009, Spain

Abstract

Gene therapy using adeno-associated virus (AAV) vectors has emerged as a promising strategy, offering new hope for patients with genetic disorders. However, during rAAV manufacturing, residual host-cell DNA (hcDNA) and plasmid DNA from the production process may remain in the final product. These impurities are classified as Critical Quality Attributes (CQAs) for release testing, and accurate quantification of these contaminants is essential to mitigate potential safety risks and ensure regulatory compliance. To address this requirement, a droplet digital PCR (ddPCR) method was developed and validated in accordance with ICH Q2(R1) guidelines for the quantification of residual E1A gene sequence, a viral oncogene constitutively expressed in Pro10TM and other HEK293-derived producer cell lines, which plays a critical role in rAAV production.

For the E1A-ddPCR method, rAAV samples were pretreated with and without DNase I, DNA was extracted using a semi-automated magnetic beads-based system and analyzed by ddPCR targeting several E1A amplicons of different lengths. This design enabled quantification of both total residual E1A DNA (free, encapsidated, or capsid-associated) and encapsidated and/or capsid-associated residual E1A DNA (DNase-resistant), as well as fragmentation analysis to assess DNA integrity.

The validated method demonstrated adequate repeatability (CV ≤ 3%), intermediate precision (CV ≤ 5%), and accuracy (83 - 100%) across a defined range using the QX200 and QX ONE ddPCR systems. Specificity against the target sequence or other components and/or impurities regularly found in the test samples was also confirmed, and robustness testing ensured reliability under variation of key parameters: different primers/probe lots, plate storage, DNase activity, and the use of QX200 vs. QX ONE system. Additionally, correlation analysis with total hcDNA levels in rAAV samples suggested similar behavior of total hcDNA based on 18S and E1A sequences during encapsidation and purification.

Overall, this analytical approach for impurity detection strengthens quality control for product characterization and process development.

Methods

E1A-ddPCR method: Samples are analyzed following the 3 main steps as follows:

DNA Extraction and Purification using a Semi-automated Magnetic Bead–Based System: The test sample is extracted following two protocols:

▪ Protocol A (DNase I): Sample treated with DNase I, followed by extraction using a commercial kit for the extraction and purification of residual DNA (semi-automated magnetic bead–based system) to quantify DNase-resistant E1A (encapsidated or capsid-associated).

▪ Protocol B (Mock): Sample treated with DNase I digestion buffer without enzyme, followed by extraction using a commercial kit for the extraction and purification of residual DNA (semi-automated magnetic bead–based system) to quantify total residual E1A (free/non-encapsidated, capsid-associated, and encapsidated).

Post-extraction Sample Dilution: Prior to ddPCR plate preparation, the extracted DNA is diluted 2-fold and 5-fold in dilution buffer at room temperature.

Quantification - Droplet Reading: After droplet generation and PCR, the instrument reader (QX200/QX ONE) analyzes each droplet individually. Positive droplets, which contain at least one copy of the target gene fragment, emit increased fluorescence compared to negative droplets.

Data Analysis: Data analysis is performed using the QX Manager Regulatory Edition software, which measures the number of positive and negative droplets in each sample, and determines the sample concentration in each well after Poisson's statistical processing, to determine the target concentration in copies per µL (c/µL). These values are multiplied by the dilution factors applied to the sample during the extraction, dilution and PCR reaction setup. To obtain final concentrations, Grubbs´ statistics is followed to detect and discard outliers within replicates, when applicable. Final results are reported in copies/mL.

The E1A-ddPCR method was developed at Viralgen Vector Core. During the validation procedure, the characteristics corresponding to a quantitative test for impurities as per ICH Q2 (R1) were assessed: specificity, accuracy, precision (repeatability and intermediate precision), linearity and range, limit of quantification, as well as robustness of the method.

Briefly,

Three primer/probe (PnP) sets targeting E1A were aligned in silico against one E1A-containing plasmid and nine process-related plasmids (GOI, Rep/Cap or Helper) lacking this gene using SnapGene software, and the human genome via UCSC In-silico PCR tool. In vitro specificity was confirmed by analyzing 14 samples: 3 containing the E1A gene and 11 lacking it.

Potential matrix interference with DNase I activity and method accuracy was evaluated through spike recovery in a representative set of final formulation buffers. Eight formulation buffers were spiked with a known amount of E1A-containing plasmid, and recoveries were calculated for each amplicon.

Method linearity was assessed in four independent assays performed by two analysts. The E1A-containing plasmid was diluted in dilution buffer and, after DNA extraction, further diluted fifteen times to cover a theoretical working range of 4–4706 copies/µL.

To assess sample linearity, accuracy, repeatability and intermediate precision, six independent assays were performed by two analysts using representative material (R-289A rAAV sample spiked with E1A-containing plasmid) at six concentrations under Mock and DNase I conditions. Accuracy was evaluated as recovery across concentrations within the defined range for the sample. Intermediate precision was also assessed using the QX ONE system.

Assay Principle

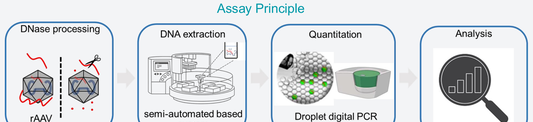

Fig 1. Graphical representation of the overlapping amplicons, and the workflow of the method. The imaged E1A-ddPCR method is comprised of an initial DNA extraction process (total and encapsidated/DNase resistant E1A), followed by sample dilution, ddPCR-mediated detection, and quantification of the target E1A sequences.

Results

Fragment Distribution Calculations

Fig 2. Example of the schematic representation of primers/probe sequences and fragment distribution calculation. Different E1A fragments are detected, allowing evaluation of the fragment size distribution in the sample. The PnP set for the Short Amplicon is considered as representative of total E1A DNA as this amplicon will also quantify Medium and Large amplicons.

Specificity

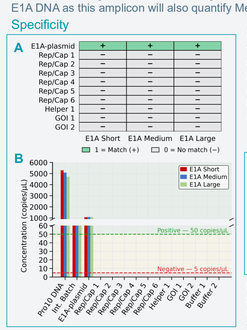

Fig 3. Specificity results for the E1A-ddPCR quantification method, both (A) in silico and (B) in vitro, on samples containing the residual E1A gene sequence. (A) Primer specificity analysis for different E1A amplicons across ten plasmid sequences. Green cells (+) indicate primer hybridization (match); light gray cells (–) indicate no primer hybridization (no match). (B) Bars represent observed concentrations for each amplicon (short (< 100 bp): red; medium (>200 and <300 bp): blue; Large (> 500 bp): green). Shaded zones indicate decision criteria: red (<5 copies/µL, negative) and green (≥50 copies/µL, positive).

Robustness

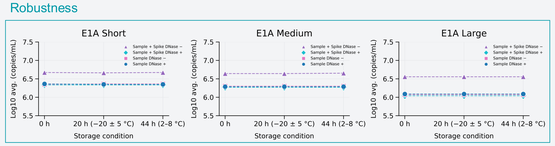

Fig 4. E1A quantification under different storage conditions. Log_{10} average concentration (copies/mL) of the E1A after different storage conditions. Data points represent individual measurements. Lines connect identical sample conditions across storage times. When identical values occurred within a condition, points were minimally offset vertically for visualization purposes only (±0.0075 log units).

Total HCDNA Correlation

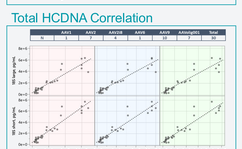

Fig 5. Correlation between values of residual E1A and total residual hcDNA based on 18S rRNA locus qPCR (Andre et al., 2016). The method was applied to a total of 30 rAAV samples from different serotypes, including AAV1 (n = 1), AAV2 (n = 7), AAV2i8 (n = 4), AAV6 (n = 1), AAV9 (n = 10), and AAVolig001 (n = 7), to characterize residual DNA content. Kendall's rank correlation coefficient (τ) was calculated for each pair of variables. The analysis revealed a strong and positive monotonic relationship between residual E1A and total hcDNA levels, with τ values ranging from 0.75 to 0.77 ($p < 0.001$), indicating that higher hcDNA concentrations are associated with a higher number of E1A gene copies. These results suggest a similar behavior of total hcDNA based on 18S rRNA and E1A sequences during encapsidation and purification, to a comparable extent across the different serotypes examined.

Conclusions

The validated ddPCR method provides accurate and precise quantification of residual E1A in rAAV products, fulfilling ICH Q2(R1) requirements for specificity, linearity and range, accuracy, precision, LOQ, and robustness. Reproducibility and applicability across different AAV serotypes in routine batch characterization confirm it's a reliable platform tool for ensuring global regulatory compliance in gene therapy products, while supporting complementary monitoring strategies for process-related impurities.

References

(1) ICH (International council for harmonization) Q2(R1): Validation of analytical procedures: text and methodology. (2) GMP Part IV. Guidelines on Good Manufacturing Practice specific to Advanced Therapy Medicinal Products. European Commission, 2017. (3) Andre M., Reghin S., Boussard E., Lempereur L., Maisonneuve S. 2016. Biologicals 44 (3) 139-149.

Validation Summary Table

Parameter	Acceptance criteria	Results
Specificity In silico	Primers and probes targeting the E1A sequence must show a 100% alignment when tested against the plasmid containing the E1A sequence and must lead to an amplicon of the expected size. On the contrary, primers and probes targeting the E1A sequence must not show any alignment leading to an amplicon formation when tested against plasmids lacking the E1A sequence or against the human genome.	- Plasmids harboring the E1A gene: 100% alignment detected and leading to an expected amplicon for the different E1A primers and probe sets - Plasmids lacking the E1A gene: No alignment detected leading to an amplicon for the different E1A primers and probe sets.
Specificity In vitro	For those samples harboring the E1A gene, a clear positive signal must be detected in the ddPCR reaction. Concentration of positive samples must be ≥ 50 copies/µL. For those samples that lack the E1A gene, results must be negative, with target signal ≤ 5 copies/µL.	-Samples lacking the E1A gene: Negative for the different E1A primers and probe set targets, ≤ 5 copies/µL. - Samples harboring the E1A gene: Positive for the different E1A primers and probe set targets, ≥ 50 copies/µL.
Specificity - Matrix interference	Lack of interference with DNase I activity will be confirmed if spike recoveries in the DNase- treated samples are ≤ 10%. Lack of interference in the quantification of residual E1A will be confirmed if the spike recoveries in the Mock condition are within the 50-150% range.	Recoveries between 96 % and 104 % for Mock treated samples. No recovery for DNase- treated samples.
Method linearity	R^2 must be ≥ 0.990 for the range of consecutive dilution levels to be included in the linear range, with Y-intercept values ≤ 1/3 of the lowest concentration considered within the linear range.	R^2 = 1.000 and Y-intercept values ≤ 1/3 for the dilution levels corresponding to the following ranges: 18 – 2676 copies/µL for the Short amplicon. 18 – 2673 copies/µL for the Medium amplicon. 11 – 2673 copies/µL for the Large amplicon.
Sample linearity	For each PnP set tested, R^2 ≥ 0.990 for the consecutive sample concentrations considered linear.	For the QX200, R^2 = 1.000 for samples within the defined range. For the QXONE, R^2 ≥ 0.999 for samples within the defined range.
Accuracy	For each PnP set, recoveries must be in the 50%-150% range.	Recoveries within 83% to 100%.
Precision – Repeatability	The CV of the residual E1A levels reported for DNase + and DNase – conditions and for each PnP set must be ≤ 30%.	CV ≤ 3%
Precision – Intermediate precision	CV values must be ≤ 40% for each E1A level tested.	CV ≤ 5%
Limit of quantification (LOQ)	LOQ is defined as the lowest E1A concentration within the sample linear range that shows an accuracy within 50-150% and precision with CV values ≤ 40%. Range of the method is defined as the interval from lowest to highest consecutive E1A levels that meet linearity, accuracy and intermediate precision criteria.	The lowest E1A concentration within the sample linear range is quantified with nominal concentration within 83 % and 115% accuracy and CV ≤ 11% (The most restrictive parameters between QX200 and QXONE).
Range	Interval from lowest to highest consecutive E1A concentrations that meet linearity (R^2 ≥ 0.990), accuracy (nominal concentration ± 50%) and precision (CV ≤ 40%) criteria.	Defined range shows R^2 ≥ 0.999, accuracy of nominal E1A concentrations ± 17% and CV ≤ 11 % for intermediate precision (The most restrictive parameters between QX200 and QXONE).
Robustness – Storage	Recoveries after storage ≥ 50% compared to the concentrations obtained on day 1.	Recoveries after storage of the extraction plates at - 20 ± 5 ºC for 20 hours were between 92% and 105%. Recoveries after storage of the dilution plates at 2-8 ºC for 44 hours were between 92% and 109%.
Robustness – PnP lots	CV ≤ 30% per primer/probe set.	CV ≤ 5%
Robustness – DNase activity	Spike clearance ≥ 90%. (Recovery ≤ 10%)	Spike recovery = 0%
Robustness – QX200-QX ONE	R^2 ≥ 0.990 for the range of consecutive dilution levels to be included in the linear range, with Y-intercept values ≤ 1/3 of the lowest concentration considered within the linear range.	R^2 = 1.000 with Y-intercept values ≤ 1/3 of the lowest concentration considered within the linear range. Same working range as for QX200.
	R^2 ≥ 0.990 for the consecutive sample concentrations considered linear in the QX200 system, recoveries of 100 ± 50% for the sample concentrations considered accurate in the QX200 system, and intermediate precision of CV ≤ 40% for the sample concentrations considered precise in the QX200 system	R^2 ≥ 0.999. Recoveries within 94% to 115%. CV ≤ 11%. Same sample linear range as for QX200.