scientific reports

Check for updates

OPEN

Three-dimensional linkage analysis with digital PCR for genome integrity and identity of recombinant adeno-associated virus

Tam Duong[1,3]✉, Michele Firmo[2,3], Chien-Ting Li[1], Bingnan Gu[1]✉ & Peng Wang[1]✉

Recombinant adeno-associated virus (rAAV) has emerged as the vector of choice for in vivo gene delivery, with numerous clinical trials underway for the treatment of various human diseases. Utilizing rAAV in gene therapy requires a highly precise quantification method to determine the viral genome titer and further establish the optimal therapeutic dosage for a rAAV product. The conventional single-channel droplet digital PCR (1D ddPCR) method offers only partial information regarding the viral vector genome titer, lacking insights into its integrity. In our pursuit of further advancing rAAV analysis, we have developed a novel 3D ddPCR assay with advanced 3D linkage analysis. We have designed the three amplicon sites targeting both ends of the viral genome, as well as the center of key therapeutic gene of interest (GOI). This study aims to offer a more comprehensive and insightful assessment of rAAV products which includes not only quantity of viral genome titer but also the quality, distinguishing between partial ones and intact full-length viral genomes with the right GOI. Importantly, due to the random partitioning property of a digital PCR system, the 3D linkage analysis of rAAV viral genome requires a proper mathematical model to identify the true linked DNA molecules (full-length/intact DNA) from the population of false/unlinked DNA molecules (fragmented/partial DNA). We therefore have developed an AAV 3D linkage analysis workflow to characterize genomic integrity and intact titer for rAAV gene therapy products. In this study, we focus on evaluating our 3D linkage mathematical model by performing DNA mixing experiments and a case study using multiple rAAV samples. Particularly, we rigorously tested our algorithms by conducting experiments involving the mixing of seven DNA fragments to represent various AAV viral genome populations, including 3 single partials, 3 double partials, and 1 full-length genomes. Across all 37 tested scenarios, we validated the accuracy of our workflow's output for the percentages of 3D linkage by comparing to the known percentages of input DNA. Consequently, our comprehensive AAV analytical package not only offers insights into viral genome titer but also provides valuable information on its integrity and identity. This cost-effective approach, akin to the setup of traditional 1D or 2D dPCR, holds the potential to advance the application of rAAV in cell and gene therapy for the treatment of human diseases.

Keywords AAV, Titer, 3D dPCR, Genome integrity, GOI identity, 3D linkage analysis

Adeno-associated virus (AAV) stands as a prominent gene delivery platform for the treatment of various human diseases. Recent clinical trials have shown the remarkable potential of AAV in addressing conditions such as aromatic L-amino acid decarboxylase (AADC) deficiency, Parkinson's disease, achromatopsia, and hemophilia B (reviewed by[1])AAV is a non-enveloped virus that can be genetically engineered into recombinant AAV (rAAV) to serve as a vector for delivering DNA to target cells. This AAV delivery method has numerous advantages, including biosafety, minimal immunogenicity, a broad range of infectivity, and the ability to sustain stable

[1]Research & Development, Lonza Houston, Inc., 14905 Kirby Dr, Houston, TX 77047, USA. [2]Global Biologics Technical Development, Lonza Basel, Inc., Muenchensteinerstrasse 38, 4052 Basel, Switzerland. [3]These authors contributed equally: Tam Duong and Michele Firmo. ✉email: tam.duong@lonza.com; bingnan.gu@lonza.com; peng.wang@lonza.com

long-term gene expression in vivo[1,2]. Recombinant AAV, however, presents a challenging drug product for characterization, primarily due to its complex biological nature. A rAAV contains a packaged viral genome that delivers the therapeutic gene of interest, often exhibiting a mixture of full, empty, and partial DNA fragments[3]. Additionally, the infectivity of a rAAV product can vary, ranging from strong to weak, in both in vitro cell culture systems and in vivo tissue-specific delivery scenarios. To gain a comprehensive understanding of a rAAV product, continuous improvement and optimization of rAAV analytical methods are essential, especially in light of the substantial growth in the application of rAAV in in vivo gene therapy[3].

In rAAV gene therapy, the process begins with the infection of rAAV into targeted cells, followed by its transportation to the nucleus through endosomes. Once inside the nucleus, the second strands of AAV are synthesized and transcribed. The translation of the therapeutic protein, encoded by the recombinant AAV viral genome carrying the gene of interest (GOI), has the potential to correct the genetic abnormalities and address the root cause of a disease (reviewed in[1,4]). Therefore, information of the viral genome is of utmost importance, encompassing aspects such as quantity (viral genome titer), quality (distinguishing partial from full-length genomes), and identity (the presence of the GOI). These factors play a pivotal role in determining the success of gene therapy using rAAV, as the viral genome titer provides crucial insights into the preclinical and clinical dosages required for an effective rAAV therapeutic product.

The primary method for quantifying the number of viral genomes is PCR. Within PCR-based approaches, two widely used analytical methods exist for determining the physical titer of AAV vector samples: traditional quantitative PCR (qPCR) and digital-based PCR methods. Real-time quantitative PCR has been extensively used for quantifying the number of DNA targets, relying on a standard curve. Digital PCR represents a next-generation technology that allows highly precise and absolute quantification of nucleic acids without the necessity of a standard curve[5,6]. The digital-based PCR methods are based on the concept of partitioning the PCR reaction into numerous small-volume compartments with nanoliter precision, where molecules are randomly distributed. There are two main systems available for partitioning the PCR reaction: one that divides it into droplets (droplet digital PCR—ddPCR) and another that partitions it into nanoplates (nanoplate digital PCR—dPCR)[7].

In the context of this study, we have developed an AAV titer assay utilizing the ddPCR system. ddPCR employs a microfluidic device to partition a PCR reaction into approximately 20,000 individual droplets, each of which functions as an independent and individual PCR reaction. The fluorescence of each droplet is measured after PCR and categorized into populations of either positive or negative droplets based on a predefined experimental threshold value. Through a combination of limiting dilution and Poisson distribution statistics, ddPCR furnishes an absolute measurement of concentrations, eliminating the need for a standard curve[8]. When applied to AAV viral genome tittering and PCR end-point measurement in AAV infectivity assay, this method has demonstrated improvements in accuracy, intra- and inter-assay precision, and increased resistance to PCR inhibitors[6,9].

Currently, the quantification of viral genome titer is achieved using ddPCR, which typically targets a single site within the viral genome. However, this one-dimensional ddPCR assay can only provide quantification for a specific segment of the viral genome. Consequently, it cannot offer insights into whether the quantified genomes are full-length or partial. The packaging of the AAV genome into viral capsids is an imperfect process, leading to the formation of partial genomes in the system. The delivery of a partial viral genome can significantly impact the success of gene therapy due to the potential missing of crucial therapeutic gene components. To address this limitation, two-dimensional ddPCR has been developed and applied for quantifying full-length AAV viral genomes[10]. In this two-dimensional assay, both ends of the viral genome are targeted and amplified, enabling the calculation of linkage between the two targets. This approach provides additional insights into viral genome integrity but still lacks information about the middle portion of the viral genome, which contains the therapeutic gene of interest (GOI). To overcome this, we have developed a three-dimensional ddPCR assay that targets three positions within the viral genome: the 5' end promoter region, the GOI, and the 3' end Poly(A) region. This advanced multi-dimensional approach aims to provide a more comprehensive view of a recombinant viral genome, allowing us to ascertain the presence of an intact viral genome with the correct GOI target linked together with correct sequence.

As the number of targets in a PCR reaction increases, there is a need to develop advanced analytical methods to enable complete three-dimensional analysis, particularly emphasizing the linkage analysis of all three targets. Multiple studies previously published describe the application of triplex digital or multiplex droplet digital PCR assays to evaluate the intact proviral HIV-1 DNA[11–13]. However, these studies did not focus on addressing the real linkage of the targets where all three targets are linked in a single DNA molecule. In digital PCR setting, it is straight forward to set up multiplex 2D or 3D dPCR reaction, in fact, we can set up as much as 5, 6 targets in one reaction. The key challenge is how to calculate the concentration of the DNA template, including all the target regions in single molecules, i.e. the level of linkage of all DNA amplicons. Due to the random distribution of multiple DNA molecules in each partition, the triple-positive cluster is not equal to triple-linked DNA molecules which represent for the real intact viral genome. When fragments of un-intact DNA are co-partitioned into one droplet, there will be formation of false triple positive droplets.

The linkage analysis for 2D dPCR was previously described with an equation to calculate true linked/ intact genome with 2 targets[14], this study described well the concept of linked and unliked DNA molecules in a 2D dPCR setting. However, there are limitations with the current 2D linkage analysis when applied to a 3D dPCR assay. Specifically, 2D linkage analysis can only provide information about pairs of targets, lacking the ability to establish a complete 3D relationship between them.

This paper is to evaluate our 3D linkage mathematical model in rAAV characterization, that can provide a comprehensive view of the viral genome including titer, integrity and identity GOI information. Through the analysis of all false-positive cases involving triple-positive droplets, we have successfully calculated the titer and percentage of 3D full-length linkage populations in an AAV viral genome product. Importantly, our model has been evaluated extensively by wet lab testing using plasmid DNA and real rAAV samples.

In summary, we have developed an advanced 3D ddPCR assay for the analysis of AAV viral genomes. This assay enables us to gain insight into the integrity of AAV viral genomes while simultaneously confirming the presence of the correct gene of interest (GOI) in a single ddPCR reaction. This innovative method represents a significant advancement in the field of AAV gene therapy, offering a more comprehensive understanding and highly accurate measurement of true intact viral genome products.

Materials and methods

HEK293 suspension cell culture

The proprietary 5B8 HEK293 (Human Embryonic Kidney Cells) suspension cells were developed from HEK293 cells from ATCC (CRL-1573.3)[15]. Cells were cultured in FreeStyle F17 medium (Gibco, A1383501), supplemented with 4 mM GlutaMAX, 0.2% Pluronic F-68. Cells were passaged at 6E + 05 vc/mL every 3 to 4 days and cultured on shakers at 120 rpm in a 37 °C humidified incubator at 5% CO_2.

HT1080 cell culture

Human epithelial cells derived from connective tissue, HT1080 cells (ATCC, Cat. CCL-121) were cultured in DMEM high glucose (HG), GlutaMAX supplement media (Gibco, 10-566-024) supplemented with 10% defined fetal bovine serum (Cytiva Hyclone, SH30070.03HI). Cells were maintained in a 37 °C humidified incubator at 5% CO_2. Cells were used from passage 3 to passage 20 for the AAV infectivity assays.

Triple transfection for AAV viral production

HEK293 suspension cells were inoculated 20 to 24 h before transfection at 1.4E + 06 vc/mL. The plasmids used for the transfection were pHelper, pAAV-GFP and pRep/Cap2, pRep/Cap8 or pRep/Cap9 for AAV2, AAV8 or AAV9 serotype respectively. AAV9 viral vector was used in 2D and 3D ddPCR experiments (Figs. 1, 2, and 3). AAV2 and AAV8 were used in 3D ddPCR assays and Nanopore sequencing experiments (Fig. 4). Briefly, the plasmids, FectoVIR are diluted separately in Opti-MEM medium (Life Technologies, Grand Island, NY, USA) and then combined. After 30 min incubation of FectoVIR and plasmids mix, the complex was added to the cell culture for transfection. The culture was continued for three days before harvest.

Suspension cultured cells were lysed using NaCl, Triton X-100, Benzonase and $MgCl_2$. The lysed cell suspension was centrifuged at 10,000 g at room temperature for 10 min, and the supernatant was collected for the later step of ddPCR analysis. rAAV from triple transfection culture was harvested and purified following downstream Lonza platform process[15].

Vectors

Plasmids used in this study were generated by Genscript. The list of plasmids representing for 7 different AAV viral vector fragments: CMV:GFP: Poly(A) (full-length); CMV: GFP, GFP:Poly(A), CMV:Poly(A) (double partial); CMV, GFP, Poly(A) (single partial). Plasmids were prepared in 1xTE buffer in industrial grade, using Stbl3 strain. The list of plasmids and their concentration is presented in Table 1 and Supplementary Table 2.

Primers and probes

Primers and probes for each target were designed following ddPCR Bio-Rad guidelines. We selected the amplicon size less than 100 bp and have a similar size range between 3 targets (CMV, GFP, Poly(A)) (Supplementary Table 1). CMV probe was tagged with VIC (channel 2), Poly(A) probe was tagged with FAM (channel 1), and GFP probe was tagged with Cy5.5 (channel 4). The ITR primers and probe were selected as previously published[10].

Sample preparation before ddPCR

The collected rAAV sample was treated with DNaseI (Promega, M610A) at 37 °C, for 1 h, then moved to the proteinase K treatment step. The samples were incubated with Proteinase K at 37 °C for 1 h, followed by 95 °C for 10 min. The sample post ProK treatment can be subjected to ddPCR run, the sample is diluted in nuclease-free water (Life Technologies, AM9935) to the concentration range from 50 – 5000 copies/uL to get in the limit of quantification range of the ddPCR system.

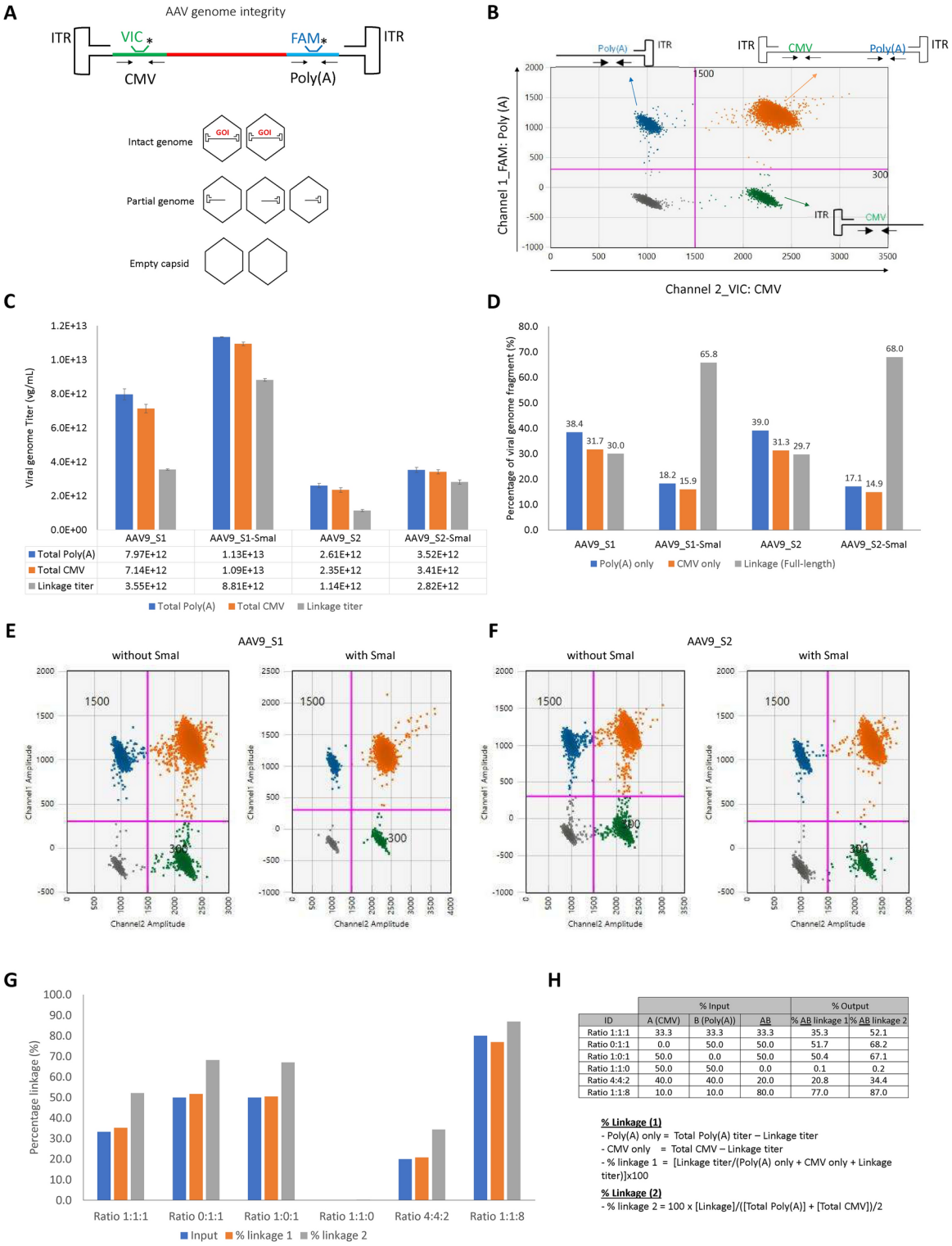
Droplet digital PCR

1D, 2D and 3D droplet digital PCR was performed in 22 µL reactions containing 5 µL of DNA sample, nuclease-free water, 3 sets of primers and probes (targeting CMV, GFP and Poly(A)), 2X ddPCR Supermix for Probes (Bio-Rad, 1,863,010). The primers were used at a final concentration of 900 nM, and the probes were used at a final concentration of 62.5 nM in a 3D ddPCR or 1D ddPCR reaction. Droplet digital PCR was performed using Bio-Rad QX ONE^{TM} system (QX ONE Software 1.2 Standard Edition) following the manufacture Bio-Rad protocol. In this system, samples were partitioned into approximately 20,000 nanoliter-sized droplets. PCR amplification program was set as: Stage 1 (95 °C, 10 min); Stage 2 (94 °C, 30 sec; 56 °C, 1 min; repeat for 40 cycles); Stage 3 (98 °C, 10 min); with temperature ramp at 2 °C/sec.

2D linkage analysis

2D linkage analysis of 2 targets is calculated based on the formula below, as previously described[14]:

$$\text{linkage} = \left[\ln(N) - \ln\left(n_{00} + n_{01} + n_{10} + \frac{n_{01}n_{10}}{n_{00}}\right)\right] * \left(\frac{1000}{\text{droplet volume}}\right)$$



where: Linkage unit = copies/μL; N = total number of droplets; n00 = number of double negatives; n01 = number of channel 1 negative and channel 2 positive; n10 = number of channel 1 positive and channel 2 negative.

The percentage linkage of a 2D ddPCR assay in this study was calculated using the formula below:

$$\text{Poly(A) only} = \text{Total Poly(A) titer} - \text{Linkage titer}$$

$$\text{CMV only} = \text{Total CMV titer} - \text{Linkage titer}$$

$$\%\ \text{Linkage} = [\text{Linkage titer}/(\text{Poly(A) only} + \text{CMV only} + \text{Linkage titer})] \times 100$$

◀**Fig. 1**. 2D ddPCR assay provides information on intact viral genome titer. (**A**) Two-dimensional targets designed for characterization of AAV viral genome integrity: CMV promoter region tagged with VIC, and Poly(A) region tagged with FAM. (**B**) The 2D Amplitude graph displays 4 populations of double positive droplets, single positive droplets and negative droplets. The subset of double positive droplets for two target sites contain a full-length intact viral genome population; the single positive droplets contain a partial viral genome population. (**C**) 2D ddPCR assay on AAV9 sample 1 (AAV9_S1) and AAV9 sample 2 (AAV9_S2). This assay provides three titer values, consisting of total Poly(A) titer, total CMV titer and linkage (full-length) titer (vg/mL). 2D ddPCR assays were performed in the presence of SmaI or without SmaI in the PCR reaction mix. (**D**) Percentage linkage and percentage of partial viral genome (Poly(A) only; CMV only) were calculated using the formula in Material and Method (2D linkage analysis section). (**E**) 2D amplitude graphs of sample AAV9_S1, with and without SmaI treatment. (**F**) 2D amplitude graphs of sample AAV9_S2, with and without SmaI treatment. With SmaI treatment, the 4 populations were better separated with less droplet rainy effect between sections. (**G**) Percentage linkage calculation in a DNA mixing experiment. Plasmid DNA contain each targets (CMV only, Poly(A) only) and both two targets (CMV:Poly(A)) were mixed in different ratio as: 1:1:1; 0:1:1; 1:0:1; 1:1:0; 4:4:2; 1:1:8. (**H**) Table shows the comparison between the percentage of input and output with two different percentage calculation formula named % Linkage (1) and % Linkage (2). Notes: The data present for 2D linkage percentage in this study used the % Linkage (1) formula. The error bars in graph 2C represent for technical triplicate of either AAV9 sample 1 (AAV9_S1) of AAV9 sample 2 (AAV9_S2).

3D linkage analysis workflow

To keep the evaluation of our mathematical model easy to follow, we simplified target names of CMV, GFP, Poly(A) as A, B, C respectively. The actual linked DNA fragment is identified with the *underline* sign, like ABC. The group of DNA fragments that have positive signals are identified with the plus sign, like ABC+.

Due to the random partition event of the DNA fragment in a droplet digital PCR assay, population with triple positive for three targets will generally contains both true and false positive droplets ABC+ (Fig. 3A). True positive are ABC+ droplets which contain ABC linked DNA. False positive are ABC+ droplets which do not contain ABC linked DNA. To develop mathematical model for quantifying the amount of real 3D ABC-linked viral genome, it is critical to identify ABC+ false positive droplets (red, Fig. 3A) over ABC+ true positive droplet (green, Fig. 3A).

Application of the statistical principles used for 2D linkage analysis[14] cannot be directly applied for 3D linkage analysis due to the presence of more possible types of DNA fragments beside ABC linked fragment (such as double linked molecules AB, BC, AC). Lonza's mathematical model is set out to overcome this 3D linkage problem by applying a sequential list of equations specifically designed to identify particular clusters of subpopulations of interest across the cluster data of digital PCR machine outcome.

We have a process work-flow to perform 3D linkage analysis for AAV viral genome (Fig. 3B). The raw cluster data outcome from the digital PCR software is collected for further analyze using our mathematical model. This raw cluster data simply contains the count for each population; either negative, single positive (A^+, B^+, C^+), double positive for two targets (AB^+, BC^+, AC^+), or triple positive for all three targets (ABC^+) (example on Fig. 3B, left table). The algorithm then provides a weighted distribution for the concentration of each fragment. This distribution forms the baseline for retrieving summary information (mean and confidence interval) about each individual fragment (Fig. 3B, central graph) (Supplementary Fig. 3). Once the distributions are retrieved, the algorithm analyzes them individually to retrieve concentration mean of each fragment and its confidence interval (Fig. 3B, right graph).

Mantis automation liquid dispensing

DNA fragments mixing experiment were performed using the Mantis automation liquid dispensing system. Low-volume chips were used to dispense the DNA fragments mixing for the extended list of 39 scenarios (Supplementary Table 3). We included two repeated scenarios (ID 31, 39) as a manual pipetting control. The liquid automation dispensing was performed following Mantis instruction protocol. The mixed DNA samples were then applied directly to the 3D ddPCR run.

rAAV transduction assay

rAAV harvested from cell lysate was subjected to purification using resin (POROS™ CaptureSelect™ AAVX Affinity Resin, Cat. A36740) to remove the high salt and detergents before cell-based infectivity assay. Purified rAAV was then sterilized by filtration and re-titer using ITR 1D ddPCR assay. HT1080 cells were seeded on 96 well plate at 1.4E4 viable cell per well (100 µL) and the plates were incubated at 37 °C with 5% CO_2 for 4–6 h before infection. The serial dilutions of rAAVs were performed as 1000x, 2000x, 4000× and 8000× using complete media. 100 µL of diluted rAAV was added to each well of the cell-seeded plate. The plate was incubated for 3 days at 37 °C with 5% CO_2 before analyzing for GFP expression level using flow cytometry system (CytoFLEX, Beckman Coulter).

Nanopore sequencing

rAAV sequencing using Nanopore technology was published previously[16] (https://nanoporetech.com/resource-centre/workflow-aav-sequencing). Purified rAAV bulk drug substances were first treated with DNaseI to remove all unpackaged free DNA using DNaseI (Promega, M610A) at 37 °C for 1 h. Viral DNA was then extracted using PureLink™ viral RNA/DNA Mini extraction kit (Invitrogen, Cat. 12,280,050). Viral DNA from extracted rAAV was subjected to library preparations for ONT sequencing by using the Native Barcoding sequencing kit 24 V14 (SQK-NBD114.24). The prepared DNA library was loaded into a MinION Flow Cell, R10.4.1 (FLO-MIN114).

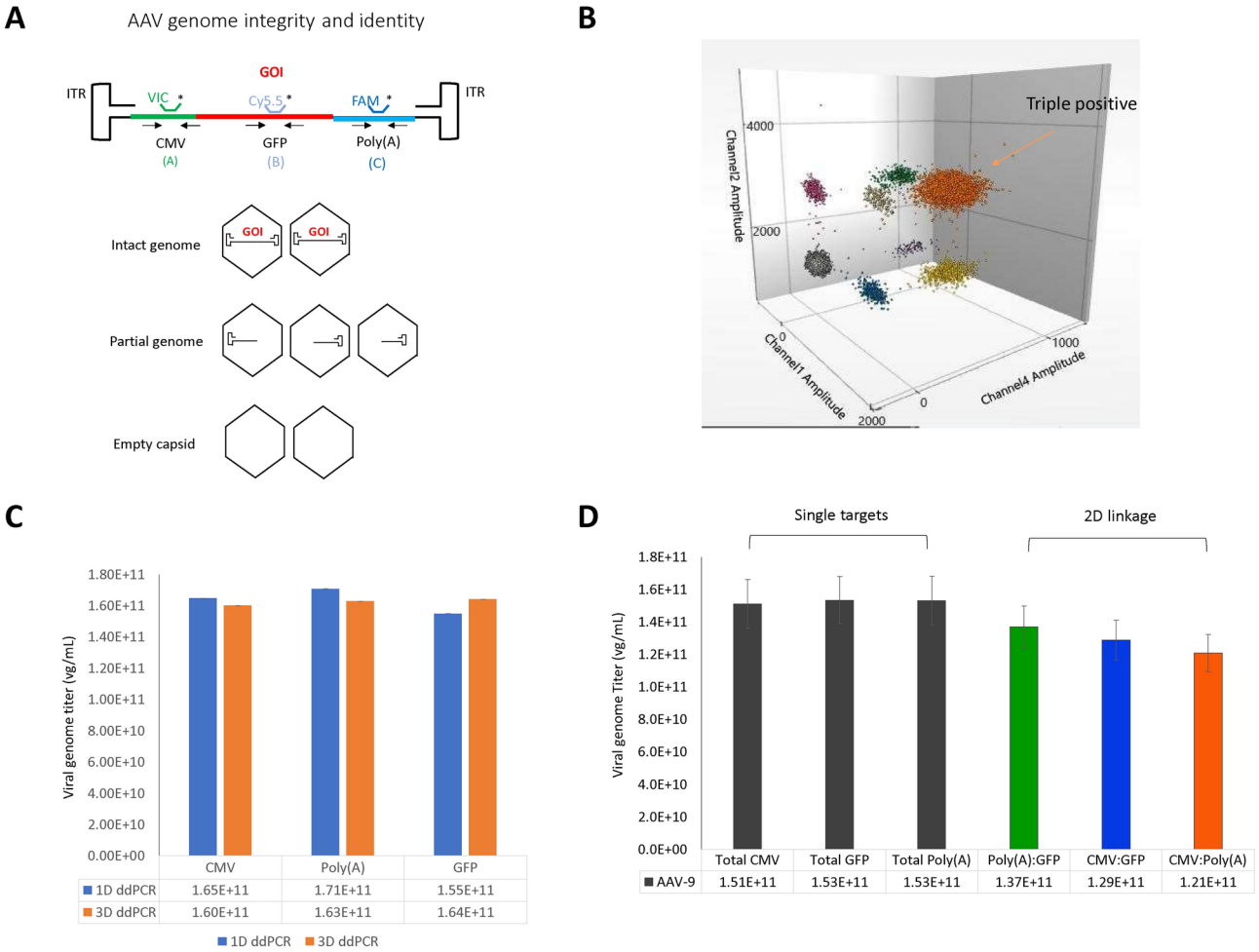


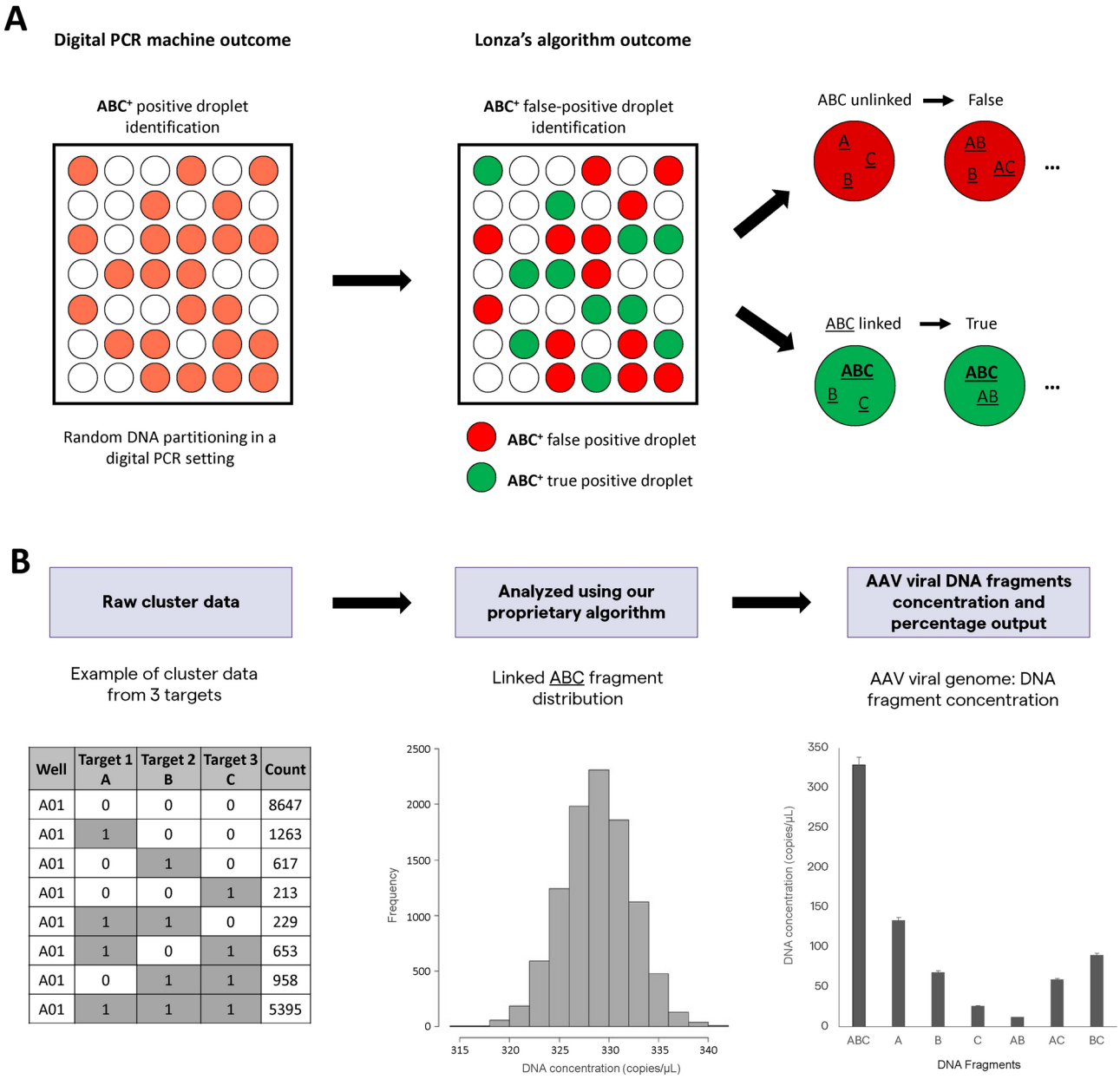
Fig. 2. Development of 3D ddPCR assay and limitation of 2D linkage analysis. (**A**) Three-dimensional targets designed for characterization of AAV viral genome integrity and identity (GOI): CMV promoter region tagged with VIC, GOI (GFP) region tagged with Cy5.5, Poly(A) region tagged with FAM. (**B**) 3D graphs show 8 populations of droplets generated in a 3D ddPCR assay where we have the orange group is the triple positive droplet population; green, yellow and white are the double positive droplet population; purple, blue and pink and purple are the single positive droplet population; grey is the negative droplet population. (**C**) Graphs show the AAV9 viral genome titer of 3 targets (CMV, Poly(A) and GFP) in 1D and 3D ddPCR assay. (**D**) Analysis of AAV9 viral genome titer from 3D ddPCR assay reveals single titers from total CMV, total GFP and total Poly(A) signals, 2D linkage titer analysis for each group of targets [Poly(A):GFP], [CMV:GFP], [CMV: Poly(A)] using Bio-Rad Quantasoft software. The error bars in graph 2C and 2D represent for technical triplicate of each viral sample.

The sequencing run was performed for 72 h using High-accuracy base calling models to produce high-quality consensus genome sequences. rAAV viral genome sequences were analyzed using the optimized EPI2ME AAV QC workflow (https://nanoporetech.com/resource-centre/workflow-aav-sequencing, https://github.com/epi2me-labs/wf-aav-qc) with additional feature analysis on full-length, single and double partial AAV fragments (Jupyter notebook and Python3). The FASTQ files were used as input into the workflow from sequencing run, along with sequence reference files for the transgene plasmid, host cell line genome, RepCap plasmid and Helper plasmid.

Results

2D ddPCR assay provides intact viral genome titer

The general droplet digital PCR work-flow is described in Supplementary Fig. 1. In a 2D ddPCR assay, we designed primers and probes targeting CMV promoter and Poly(A) regions (Fig. 1A,B). The linkage titer between these two targets provides valuable insights into the presence of intact viral genomes within the entire population. However, the accuracy of calculating the linkage of intact viral genome also depends on the maximum efficiency of PCR reactions between the two target sites. It is crucial to note that the two ends of the AAV viral genome contain two Inverted Terminal Repeat (ITR) with secondary structures that can potentially affect the accessibility of primers and probes to target sites adjacent to the ITR region.



Well	Target 1 A	Target 2 B	Target 3 C	Count
A01	0	0	0	8647
A01	1	0	0	1263
A01	0	1	0	617
A01	0	0	1	213
A01	1	1	0	229
A01	1	0	1	653
A01	0	1	1	958
A01	1	1	1	5395

Fig. 3. Development of 3D linkage analysis workflow. (**A**) Schematic represents droplet scenarios by random partition event of a droplet digital PCR assay. Orange droplets are triple positive for three targets. To develop equation/algorithm for real 3D ABC-linked viral genome, it is critical to identify ABC + false positive droplet (red, without ABC-linked DNA) over ABC + real positive droplet (green, with ABC-linked DNA). (**B**) Process work-flow to perform 3D linkage analysis for AAV viral genome: Collection of the raw cluster data outcome from the digital PCR software (Left table). After loading the data, the algorithm analyzes them providing a weighted distribution for the concentration of each fragment. This distribution forms the baseline for retrieving summary information (mean and confidence interval) about each individual fragment (central graph). Once the distributions are retrieved, the algorithm analyzes them individually to retrieve each fragment mean and confidence interval. These two pieces of information are then represented in a visually bar plot (right graph).

As previously published, the use of restriction enzyme during ddPCR can enhance the PCR efficiency and therefore provide a more accurate reflection of viral genome titers[17]. Here, we observed a similar effect of SmaI treatment on PCR efficiency. Two recognized cutting sites of SmaI restriction enzyme from each ITR can remove the secondary structures thereby improving the accessibility of primers and probes to the target sequences.

To evaluate the performance of the 2D ddPCR assay, we tested it on two purified AAV9 samples, named as AAV9_S1 and AAV9_S2. The application of SmaI treatment led to a notable increase in viral genome titers, approximately in the range of 30–40%. Importantly, there was a significant impact of SmaI treatment which led to an increase more than two times of linkage titer (Fig. 1C). This effect of SmaI treatment was significant,

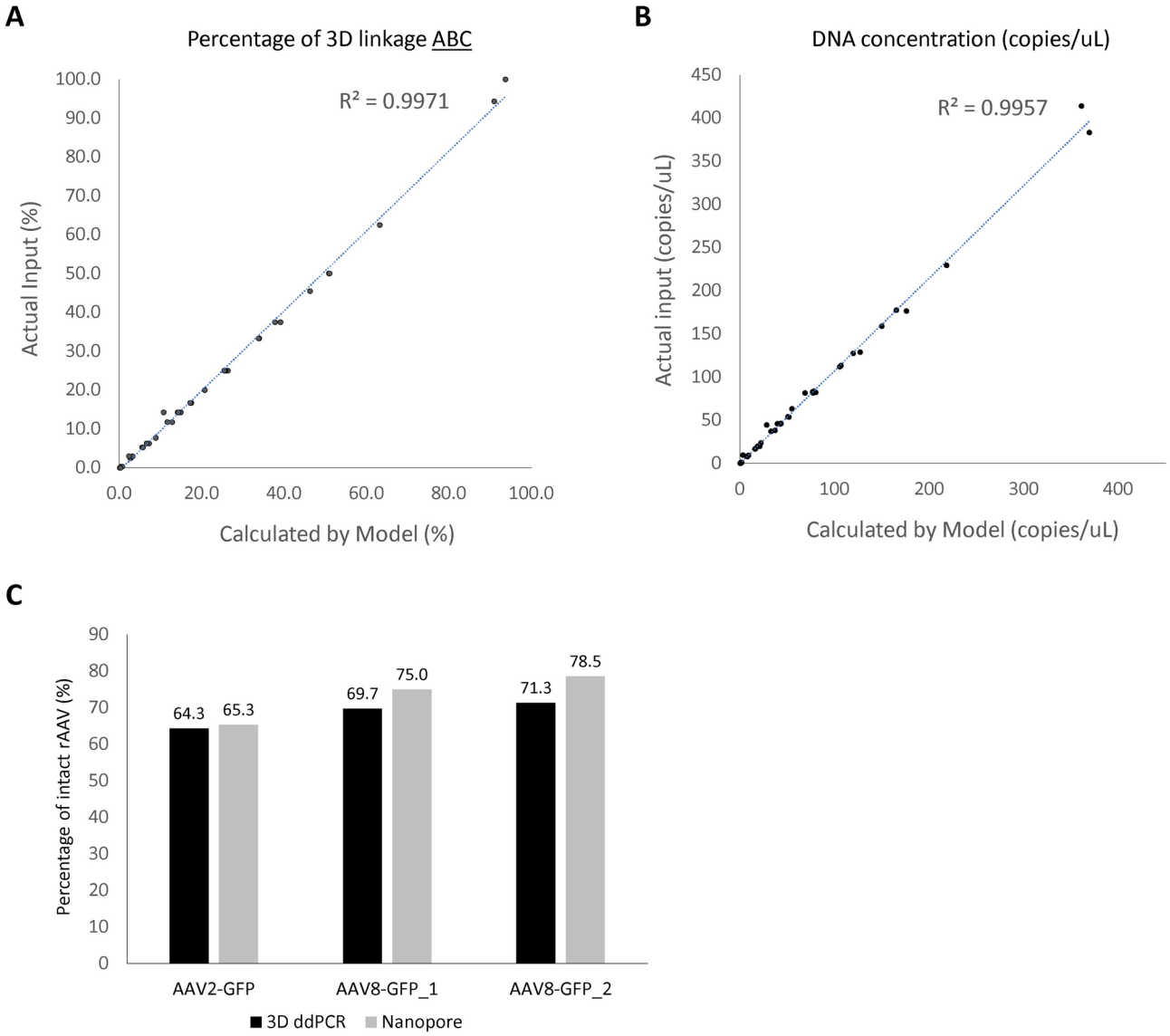
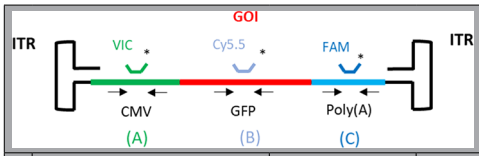


Fig. 4. 3D linkage workflow provides accurate information on real ABC linkage percentage. (**A**) The percentage of real 3D linkage ABC from the machine model output was strongly correlated to the actual input value. The data shows that the 3D linkage equation provides accurate information on actual ABC linkage percentage. (**B**) Concentration graph shows the high correlation between copies/µL of the value calculated by model and the actual input in 37 scenarios. We have 2 extra conditions which are the manual pipette control. For conditions ID 31, 39, DNA mixing was performed by using manual pipetting. For all other conditions, DNA mixing was performed using a Mantis liquid dispenser. (**C**) Intact viral genome analysis with 3D ddPCR assay and Nanopore sequencing method. Various AAV bulk drug substances (AAV2-GFP, AAV8-GFP_1 and AAV8-GFP_2) were tested for percentage of intact viral genome using 3D ddPCR assay and Nanopore sequencing analysis.

resulting in a substantial increase in the percentage of full-length viral genomes, elevating it from 30% to 65.8% in AAV9_S1 and from 29.7% to 68.0% in AAV9_S2 (Fig. 1D). In the 2D amplitude graphs, we observed a significant reduction in the "rainy effect" between populations, resulting in a cleaner separation between double-positive and single-positive population (Fig. 1E,F). This visual effect explains the improvement of PCR efficiency of the two targets.

Currently, there exists a certain degree of discrepancy in calculating the percentage of full-length genomes using the linkage titer value. To evaluate the correct formula for accurately calculating the percentage of full-length genomes in relation to partial ones, we performed DNA mixing experiments using plasmids containing CMV only, Poly(A) only and full-length viral genome having both CMV and Poly(A) regions. The DNA mixing experiment was performed at different ratios of [CMV only: Poly(A) only : CMV-Poly(A) full-length] as 1:1:1; 0:1:1; 1:0:1; 1:1:0; 4:4:2; 1:1:8 respectively (Fig. 1G), and two tested percentage linkage (full-length) formula are shown as [% Linkage (1)] and [% Linkage (2)]. The [% Linkage (2)] formula was applied as previously



	AAV viral genome sequence	Simplified name	Truncated region	Constructs
1	CMV:GFP:Poly(A)	ABC	–	
2	CMV:Poly(A)	AC	GFP	
3	GFP:Poly(A)	BC	CMV	
4	CMV:GFP	AB	Poly(A)	
5	CMV	A	GFP, Poly(A)	
6	GFP	B	CMV, Poly(A)	
7	Poly(A)	C	CMV, GFP	

Table 1. Plasmid list contains full-length or truncated forms of viral genome. To test the 3D linkage workflow, an in vitro plasmid DNA mixing conditions (37 scenarios) were performed in a DOE table with an extended list of scenarios (Supplementary Table 3). Each plasmid contains either full-length or 6 different truncation form of AAV viral genome.

published[14,17]. The mixing DNA experiment results clearly shown that [% Linkage (1)] provides a more accurate percentage which aligns closely with the actual input value when compared to [% Linkage (2)]. [%Linkage (2)] formula over-estimates the percentage of full-length viral genome population due to its omission of 50% of single populations containing either CMV or Poly(A) only. Consequently, the lower the actual AB linkage input, the more different the calculation value of [%Linkage (2)] becomes. Particularly, in the ratio 4:4:2 where we have the lower input of AB, the % output using formula (1) is 20.8% which is closely aligned with the actual input of 20%. Whereas, the % output using formula (2) is 34.4%, which is much higher and over-estimates the actual % linkage (Fig. 1G,H; Supplementary Table 7). The result suggests that the formula (1) should be utilized to accurately calculate the percentage of rAAV full-length viral genome.

3D ddPCR assay development for comprehensive rAAV viral genome characterization

To advance the analytical methods for rAAV, we have developed a novel 3D ddPCR assay designed to assess viral genome integrity and confirm the identity of the gene of interest (GOI). This advanced assay utilizes three sets of primers and probes targeting key regions, including CMV, GFP, and Poly(A) regions, as depicted in Fig. 2A. The resulting fluorescent signal data from this assay is visualized in a 3D amplitude graph, which effectively separates the outcomes into eight distinct signal populations (Fig. 2B).

To validate the effectiveness of this 3D ddPCR assay, we compared it to the traditional 1D ddPCR method. We observed similar viral genome titers for each target (CMV, Poly(A), and GFP) when using both the 1D and 3D ddPCR assays (Fig. 2C). Notably, the 1D amplitude graphs of the 1D ddPCR assay consistently exhibited clear separation between the positive and negative droplet populations in all three channels (Supplementary Fig. 2A). These results reaffirm that our target design meets expectations without any observable inhibition or competition between the targets.

In the initial stages of developing the 3D ddPCR assay, we successfully confirmed that the PCR worked precisely as expected. However, our next challenge was to analyze the resulting data in a 3D context. Unfortunately, the current software available, namely QuantaSoft by Bio-Rad, is equipped for only 2D linkage analysis. This limitation restricts the full potential of the 3D ddPCR assay, as it can only facilitate the analysis of pairs of targets at a time. Specifically, we can obtain information on single target titers such as total CMV, total GFP, and total Poly(A) titers, as well as the linkage titer of each pair of combinations, including Poly(A):GFP, CMV:GFP, and CMV:Poly(A), as depicted in Fig. 2D.

In our 2D linkage analysis, we observed higher linkage percentages between the two closer targets, such as CMV-GFP at 73.2% and GFP-Poly(A) at 80.8%, compared to the full-length linkage of CMV-Poly(A) at 65.9% (Fig. 2D, Supplementary Fig. 2B). However, these insights fall short of representing the complete complexity of the viral genome, especially when dealing with single partial fragments (CMV, Poly(A) only) and double partial fragments (CMV:GFP, GFP:Poly(A)). 2D amplitude graphs, as shown in Supplementary Fig. 2D, can only provide a 2D relationship between two targets within the 3D ddPCR assay.

To gain a more comprehensive understanding of viral genome fragments and their relationships, it is crucial to adopt 3D linkage analysis. This approach will allow us to capture the intricate interplay of all targets in the assay, providing a full and more accurate depiction of the viral genome's structure and composition.

3D linkage workflow provides accurate information on real ABC linkage percentage

The challenge in 3D ddPCR analysis is to identify false-positive droplets in the triple positive droplet population (Fig. 3A). In our 3D ddPCR assay, we targeted three regions of rAAV viral genome named as CMV(A), GFP (B) and Poly(A) (C) . The rAAV viral genomes exist in different forms as follow: all linked (ABC full-length), pairwise linked (double partial: AB, BC, AC), or not linked at all (single partial: A, B, C).

To assess the accuracy of our mathematical model and our workflow of 3D ddPCR linkage analysis (Fig. 3B; Supplementary Fig. 3), we conducted a comprehensive DNA fragment mixing experiment within the context

of a 3D ddPCR assay. In this assay, droplets could contain various combinations of DNA fragments: ABC, AB, AC, BC, A, B, C, and negative droplets (Ø). We prepared seven representative fragments for the 3D assay, encompassing ABC, AB, AC, BC, A, B, and C (as outlined in Table 1, Supplementary Table 2). Subsequently, we performed a mixing experiment that comprised 37 scenarios (as detailed in Supplementary Table 3).

To achieve precise DNA mixing in a low-volume setting characterized by the high complexity of seven fragments, we harnessed the capabilities of an automated dispenser device, Mantis (Formulatrix). Additionally, we included two control samples (ID 31 and 39) in which manual pipetting was employed for mixing, mixing ratio of ID31 is same as ID 35, and mixing ratio of ID 39 is same as ID 30 (Highlighted in grey on Supplementary Table 3). Leveraging the known copy number values of the original fragments used in this experiment (as presented in Supplementary Table 2), we calculated the original percentage input of each fragment in each scenario (Supplementary Table 4).

Our observations yielded a significant alignment of the percentage and concentration values of 3D linked ABC, between the actual input and values derived from our 3D linkage workflow, with R^2 of 0.997 and 0.996 for their correlation respectively (as depicted in Fig. 4A,B, and detailed in Supplementary Tables 4, 5, 6). This rigorous experiment, encompassing 37 diverse mixing scenarios, provides extensive evidence that our mathematical model accurately predicts the actual input of DNA fragments within a given 3D ddPCR assay.

To further evaluate our 3D linkage analysis for rAAV, we compared rAAV full-length analysis between 3D ddPCR assay and Nanopore sequencing method. We characterized the viral genome population derived from purified AAV2-GFP, AAV8-GFP_1 and AAV8-GFP_2 bulk drug substances using both methods. These rAAVs were produced using our triple transfection platform with HEK293 suspension culture system[15]. Our analysis shows that the percentage of intact rAAV is in comparable range between the two methods, 3D ddPCR and Nanopore sequencing, 64.3% vs 65.3% for AAV2-GFP; 69.7% vs 75.0% for AAV8-GFP_1 and 71.3% vs 78.5% for AAV8-GFP_2 respectively (Fig. 4C). We also observed a similar trend of percentage full-length vs single and double partials between the two methods (Supplementary Fig. 4). These results indicate that our fast and low-cost 3D ddPCR assay with linkage mathematical model can accurately calculate the concentration and percentage of full-length AAV viral genome with the right GOI identity.

Application of 3D ddPCR assay for rAAV genome titer, integrity and identity

For the application of our 3D method, we did a case study by generating rAAV2 containing multiple genome constructs through triple transfection of pRep-Cap2, pHelper and pAAV-GFP in our HEK293 suspension culture system[15]. There are 7 viral genome constructs that are either full-length (CMV-GFP-Poly(A)), or mutated forms with 100 bp truncation in each target region (CMV, GFP or Poly(A)) (Fig. 5). First, we identify the physical viral genome titer with 1D ddPCR method targeting universal ITR region. 1D ddPCR reports similar titer without genome integrity information (Fig. 5A). Our 3D linkage analysis can provide a comprehensive information of AAV viral genomes including full-length genome, double partial and single partial genomes. Only full-length AAV has the population of intact viral genome with all three targets linked (Fig. 5B). Whereas, the truncation AAV has partial viral genome with linkage of two targets (Fig. 5C–E) or no linkage detected (Fig. 5F–H) (Supplementary Table 8).

Next, we further performed transduction assay with HT1080 cell line using these rAAVs. The microscope images showed the GFP expression in infected cell of UI control, AAV full-length and truncation group (Fig. 6A). Flow cytometry analysis quantified the %GFP (+) cell population (Fig. 6B) which was further used for calculation of infectious titer (TU/mL) and specificity (vg/TU) (Supplementary Table 9 and Fig. 6C,D). We observed the correlation of transduced GFP expression with integrity of AAV viral genome. The truncation in selected CMV enhancer region minimally affected the level of GFP expression. Whereas, the Poly(A) truncation dramatically reduced GFP expression. Interestingly, the truncation in both CMV enhancer and Poly(A) regions can rescue the expression of GFP as compared to full-length control group (Fig. 6C). This case study emphasizes the value of 3D linkage method which can provide a more useful viral genome information of a rAAV product during early stages of process development.

Discussion

Recombinant AAV (rAAV) has emerged as the preferred viral vector for in vivo gene delivery owing to its exceptional biosafety, minimal immunogenicity, wide in vivo infectivity/gene delivery capabilities, and sustained expression[1,2]. As the demand for rAAV in the field of gene therapy continues to rise, it becomes imperative to constantly enhance and optimize the analytical aspects of rAAV manufacturing[3]. This is essential for the precise characterization and comprehensive understanding of the rAAV product. The viral genome titer of rAAV serves as a crucial parameter for determining the preclinical and clinical therapeutic dosage of a rAAV product. The viral genome details, encompassing aspects such as quantity (viral genome titer), quality (partial vs. full-length), and identity (Gene of Interest or GOI), are pivotal elements of a rAAV therapeutic product. To gain a thorough comprehension of a rAAV product, ongoing enhancements and optimizations of rAAV analytical methods are essential to align with the significant expansion of rAAV applications in in vivo gene therapy.

At an early stage, the qPCR was a primary method to quantify rAAV viral genome. qPCR requires a good standard material to precisely quantify DNA copy products. Digital PCR (dPCR) then becomes a revolutionary technique for precise and absolute quantification of DNA copy number without the need of a standard curve[8,18]. In digital PCR system, DNA or RNA is partitioned into numerous small-volume compartments with nanoliter-size precision, resulting in the random distribution of these molecules. Each compartment can contain zero, one or multiple molecules and the fluorescent signal of each compartment is quantified after a PCR amplification process. The initial copy number and density of target DNA are determined using Poisson statistics based on the number of PCR-positive reactions[7]. To determine the rAAV titer, a 1D dPCR assay can be designed by targeting a specific region on rAAV viral genome. This method can provide a high-sensitivity and precision for an absolute

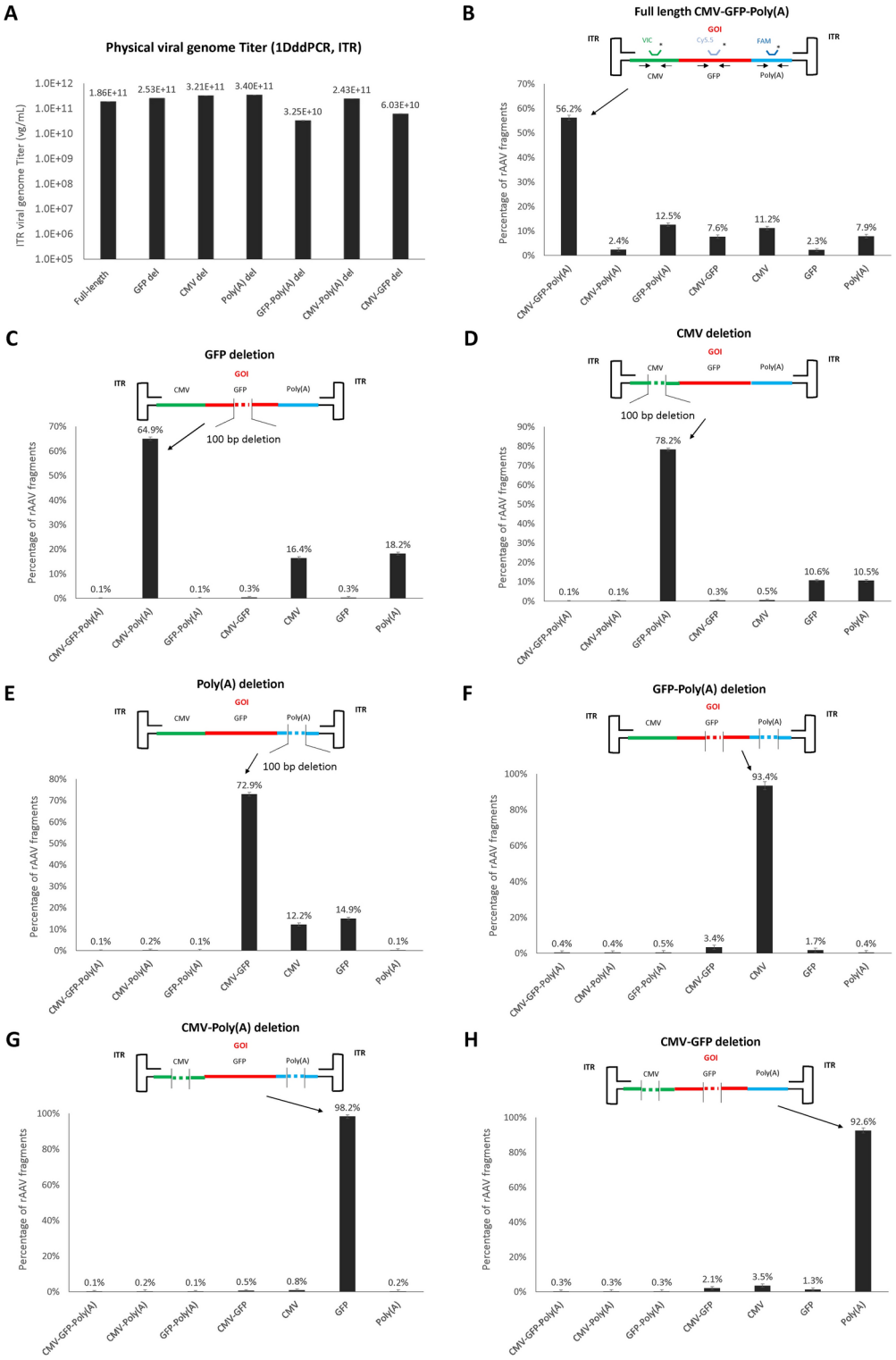


Fig. 5. Application of 3D ddPCR assay in analysis of full-length and truncated rAAV viral genome. rAAVs with full-length or truncated genome were generated by triple transfection process. (**A**) 1D ddPCR reports similar titer between full-length and truncated rAAV without genome integrity information (**B**–**H**) 3D linkage analysis provides a comprehensive picture of various rAAV population with full-length or truncated genomes. The error bars in the graph represent for technical triplicate of each viral sample.

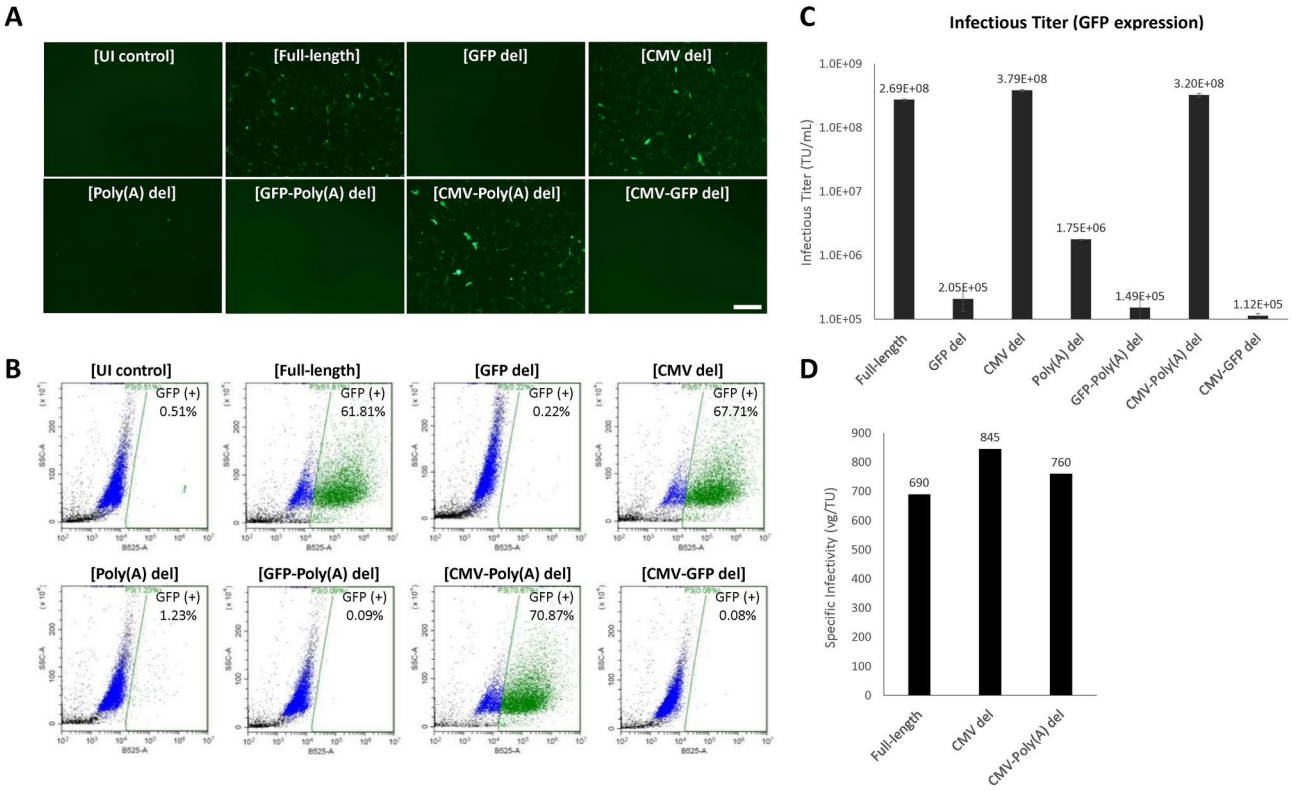


Fig. 6. Application of 3D ddPCR analysis in rAAV to correlate the product potency in vitro. rAAV infectivity assay on HT1080 cells using rAAV with full-length or truncated genome. Equal MOI was used in infectivity analysis based on conventional 1D ddPCR titer. (**A**) Microscope images of HT1080 cells day 3 post-infection with full-length and truncated AAV2-GFP. Truncation in the selected CMV enhancer region minimally affects GFP expression. (**B**) Quantitative of percentage GFP(+) cell population by flow cytometry analysis. (**C**) Infectious titer (TU/mL) was calculated based on GFP expression in infected cell population. (**D**) Specific infectivity (vg/TU) was calculate using conventional 1D ddPCR titer. Scale bar: 200 µm. MOI: Multiplicity of infection. The error bars in the graph represent for technical triplicate of each viral sample.

quantification of a rAAV product. However, AAV packaging process is not completely efficient. It also depends on the recombinant AAV design and viral genome size. Therefore a single targeting assay in 1D dPCR titer method might miss out information of partial viral genome in the whole population. 2D dPCR assay then was developed to distinguish between full-length and partial AAV products by targeting the two end regions of viral genome[10]. By choosing the right targets to detect the two-end of viral genome, 2D dPCR assay can quantify intact viral genome titer with a proper 2D linkage tool provided. This tool is currently available in Bio-rad Quantasoft software which is based on the false-positive elimination[14,17]. The false positive compartment is defined as the presence of two partial fragments containing two PCR targeting sites in one compartment. The missing component of this 2D ddPCR assay is the lack of the identity and the linkage of the therapeutic gene of interest (GOI) which is the middle part of viral genome. In some designs, one of the targets within the GOI region is selected and the other target is either at the promoter or Poly(A) region. This approach however will not provide the complete intact viral genome picture as the linkage analysis will only perform half of the viral genome length.

Our data also provide insight into the difference in % linkage (intact) depending on the target site selection. The linkage percentage of intact viral genome between CMV promoter and Poly(A) target is about 65.9%. Whereas the linkage percentage between CMV target and GOI or the Poly(A) target and GOI are higher at 73.2% and 80.8% respectively (Fig. 2D, Supplementary Fig. 2B). The length or distance between two targets is negatively correlated with the percentage linkage value. The longer the distance between two targets, the lower the linkage value will be. In fact, with our primers and probes design, the distance between CMV and GOI is around 1400 bp, whereas the distance between Poly(A) and GOI is around 800 bp. In the current gene therapy field, there is no single standardized AAV titer method.

Each laboratory or organization has its own sample preparation protocol and its own PCR reaction method. This will contribute to the different titer values reported between organizations. In our approach, we used SmaI restriction enzyme in PCR master mix previously reported as critical in digesting 2 ends of ITR to remove the secondary structure of viral genome[17]. This removal of secondary structure might increase accessibility of primers and probes to targeting sites near the ITR region. Our result showed improvement in PCR efficiency where we obtained about 30–40% higher titer targeting CMV or Poly(A) only (Fig. 1C). The percentage of

linkage is also significantly increased from 30% to 65.8% for AAV9_S1 and better reflects the true intact viral genome value as compare to non-SmaI treatment (Fig. 1D).

Further, we also carefully revisited the percentage linkage (full-length) calculation method by performing a DNA mixing experiment. We identified that the percentage linkage calculation that we used in our study reflected a more accurate percentage value as compared to formular previously published ([14,17]). The previously published formula provide overestimate the percentage linkage, and this overestimate will become significant when the portion of partial genome increases (eg. in the mixing ratio of 4:4:2 for CMV only : Poly(A) only : CMV-Poly(A) full-length) (Fig. 1G,H, Supplementary Table 7). Additionally, the published formula only provided the % of linkage population and missed out on the % of partial CMV only or Poly(A) only population. Taken together, the addition of extra details in the protocol for 2D intact titer assay with proven experimental results is very critical for the successful characterization of an rAAV product. With a strong motivation and aim to further advance rAAV analytics, we have developed a 3D ddPCR assay with an advanced 3D linkage analysis capability. This innovation aims to offer a more comprehensive and insightful assessment of an rAAV product. Using our 3D assay platform, we can identify quantity (viral genome titer), quality (distinguish between intact full-length viral genome vs. partial) and identity (information of the key GOI, therapeutic gene component). The 3D ddPCR assay is performed similarly to our current 2D ddPCR method with the addition of the third primers and probe set targeting the GOI region. However, there is a limitation in analysis of this 3D ddPCR assay using current 2D linkage analysis. Particularly, with the 2 D linkage analysis tool, we are only able to analyze the linkage target by a pair of two, either A with B, B with C and A with C. With this 2D linkage analysis, we can only provide partial information by each pair separately. Whereas, the 3D ddPCR assay should provide a more comprehensive characterization of rAAV viral genome with information on 7 species including single partial rAAV (A, B, C); double partial fragments (AB, BC, AC) and full-length rAAV with the right GOI (ABC) (Supplementary Fig. 2C).

As we understand the digital PCR technology relies on the random partitioning of DNA into separated nanosized compartments, the key to 3D linkage analysis is development of a mathematical model to identify all false positive droplets. Due to a random partition event, there will be a droplet group that has triple positive signal without a real $\underline{ABC}^{+}$ linked molecule (Fig. 3A). By identifying all of the false positive cases and building the mathematical model/algorithm to calculate this population, we can successfully calculate the percentage of full-length 3D-linked $\underline{ABC}^{\pm}$ molecules as well as the rest of the partial population (Fig. 3B). Importantly, the model is evaluated with DNA mixing samples in an actual ddPCR setting experiment. We take advantage of Mantis automation dispensing capability to perform 7 DNA fragment mixing experiments for 37 different scenarios (Supplementary Table 3). The experimental results showed a solid performance of our 3D linkage algorithm where we observed the close estimate of percentage linkage $\underline{ABC}^{+}$ and their concentration value across 37 scenarios. Our 3D linkage analysis on percentage full-length of rAAV with the right GOI was also well correlated with Nanopore sequencing analysis (Fig. 4C, Supplementary Fig. 4). This orthogonal study further strengthens the accuracy and consistency of our 3D linkage method for intact viral genome characterization. Importantly, the 3D ddPCR assay with 3D linkage analysis is a more quantitative method, providing viral genome titers of each fragment (Supplementary Fig. 4). Whereas, Nanopore sequencing provides better visualization of where the truncations or deletions occur with the viral genome. Together, these two methods can complement each other, giving us a more comprehensive understanding of the AAV gene therapy product.

In our case study, we generated rAAV with either full-length or various truncated forms of viral genome (Fig. 5). We aim to keep the viral genome size within a similar range by making a short deletion of 100 bp in each region of interest (CMV, GFP or Poly(A)) to ensure consistent comparison. However, the limitation of this case study is that the truncated viral genome, with these 100 bp deletions at predefined regions, may not accurately represent actual AAV genome scenarios. Nevertheless, this study can demonstrate the potential for detecting mutated AAV in approximate areas, depending on the amplicon targets. The traditional 1D ddPCR targeting universal site of ITR provided comparable viral genome titers between rAAV constructs without any information regarding viral genome integrity (Fig. 5A). With 3D ddPCR assay and its 3D linkage analysis, we obtained a comprehensive picture of viral genome with information of titer, integrity and identity of rAAV product (Fig. 5B–H), which was better correlated with the functional titers by the GFP expression in rAAV-GFP transduced cells (Fig. 6). This case study demonstrates the importance and value of comprehensive rAAV viral genome analysis by using 3D ddPCR method. It is important to note that in an actual AAV product, especially one with a large GOI, identifying mutation or truncation sites may be more challenging because this PCR detection method relies on the selection of amplicon target sites. To partially overcome this limitation, we can design multiple PCR target sites (eg. at the beginning, middle and end of the GOI). Combined with 3D linkage analysis, this approach may help identify potential truncation areas or recombination events in the GOI of viral genome which could better correlate with functional assay (eg. GOI expression).

For additional potential applications of this 3D linkage analysis, we can also apply it to 2D ddPCR assays which can provide closely estimated percentage linkage values compared to actual DNA-linked molecule input (Supplementary Table 7). Furthermore, our 3D linkage analysis can potentially be expanded to detect for replication-competent AAV (rcAAV) where we design different sets of primers and probes targeting wild-type AAV viral genome components, such as ITR, Rep and Cap. The 3D linkage detection of all three components will provide sufficient detection of any wild-type or rcAAV in the final product. This novel approach will advance the gene therapy field, owning its precision, sensitivity and low-cost assay as compared to next gene sequencing approach. Although our study mainly focuses on single-stranded AAV (ssAAV), it is also important to further apply the method to self-complementary AAV (scAAV), which is also increasingly used and an important class of AAV product in the field due to its proven stability and effectiveness in transgene expression[19–21].

Our current method is developed based on a Bio-rad droplet digital PCR system, it can also be further expanded into other digital PCR systems (e.g. Nanoplate digital PCR system) where the core of their technology

is also based on the molecules partitioning in precise nanoliter-sized compartments. Broader scientific community might benefit from the launching of Lonza's software application/digital portal for 3D dPCR linkage viral genome analysis in the near future.

We recognize the recently published study on intact viral genomes using the 8.5 k 96-well Nanoplate (Qiagen) and QIAcuity software analysis. In late-stage process samples, this study reported a high level of AAV sample heterogeneity, with 10–40% of the transgenes being full-length[22]. A future comparative study between the two intact viral genome analysis methods would be valuable, as it could advance the field of gene therapy by enabling more accurate and precise titer and integrity measurements of the final AAV gene therapy product.

In summary, as the field of cell and gene therapy continues to rapidly evolve and adapt to meet growing treatment demands, our objective is to advance our AAV characterization method through innovative approaches. We are expanding our methodology, from refining sample preparation to enhancing dPCR reactions and incorporating additional critical features of product into the assay development. Our goal is to offer a comprehensive AAV analysis package that not only provides insights into viral genome titer but also delivers information on its quality and identity. The rAAV quality and identity will undoubtedly be critical information to determine the success of a gene therapy product. This cost-effective approach, which is similar in setup to traditional 1D or 2D dPCR, promises to propel the application of rAAV in cell and gene therapy for the treatment of human diseases.

Data availability

The Nanopore sequencing datasets generated and/or analyzed during the current study are available in the Sequence Read Archive (SRA) repository, with Accession number of: PRJNA1111100, web link: https://www.ncbi.nlm.nih.gov/bioproject/1,111,100.

Received: 2 May 2024; Accepted: 22 October 2024
Published online: 16 January 2025

References

1. Wang, D., Tai, P. W. L. & Gao, G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* https://doi.org/10.1038/s41573-019-0012-9 (2019).
2. Pfeifer, A. & Verma, I. M. Gene therapy: Promises and problems. *Annu. Rev. Genom. Hum. Genet.* https://doi.org/10.1146/annurev.genom.2.1.177 (2001).
3. Green, E. A. & Lee, K. H. Analytical methods to characterize recombinant adeno-associated virus vectors and the benefit of standardization and reference materials. *Curr. Opin. Biotechnol.* https://doi.org/10.1016/j.copbio.2021.06.025 (2021).
4. Samulski, R. J. & Muzyczka, N. AAV-mediated gene therapy for research and therapeutic purposes. *Annu. Rev. Virol.* **1**, 427–451 (2014).
5. Whale, A. S. et al. Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res.* **40**(11), e82 (2012).
6. Lock, M., Alvira, M. R., Chen, S. J. & Wilson, J. M. Absolute determination of single-stranded and self-complementary adeno-associated viral vector genome titers by droplet digital PCR. *Hum. Gene Ther. Methods* **25**(2), 115–125 (2014).
7. Mao, X., Liu, C., Tong, H., Chen, Y. & Liu, K. Principles of digital PCR and its applications in current obstetrical and gynecological diseases. *Am. J. Transl. Res.* **11**(12), 7209 (2019).
8. Hindson, B. J. et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* **83**, 8604–8610 (2011).
9. Duong, T. et al. Improvement of precision in recombinant adeno-associated virus infectious titer assay with droplet digital PCR as an endpoint measurement. *Hum. Gene. Ther.* **34**, 742–757 (2023).
10. Furuta-Hanawa, B., Yamaguchi, T. & Uchida, E. Two-dimensional droplet digital PCR as a tool for titration and integrity evaluation of recombinant adeno-associated viral vectors. *Hum. Gene Ther. Methods* **30**(4), 127–136 (2019).
11. Levy, C. N. et al. A highly multiplexed droplet digital PCR assay to measure the intact HIV-1 proviral reservoir. *Cell Rep. Med.* **2**, 100243 (2021).
12. Levy, C. N. et al. HIV reservoir quantification by five-target multiplex droplet digital PCR. *STAR Protoc.* **2**, 100885 (2021).
13. van Snippenberg, W. et al. Triplex digital PCR assays for the quantification of intact proviral HIV-1 DNA. *Methods* **201**, 41–48 (2022).
14. Regan, J. F. et al. A rapid molecular approach for chromosomal phasing. *PLoS One* **10**, e0118270 (2015).
15. Gu, B. et al. Establishment of a scalable manufacturing platform for in-silico-derived ancestral adeno-associated virus vectors. *Cell Gene Ther. Insights* **4**, 753–769 (2018).
16. Namkung, S. et al. Direct ITR-to-ITR nanopore sequencing of AAV vector genomes. *Hum. Gene. Ther.* **33**, 1187–1196 (2022).
17. Prantner, A. & Maar, D. Genome concentration, characterization, and integrity analysis of recombinant adeno-associated viral vectors using droplet digital PCR. *PLoS One* **18**, e0280242 (2023).
18. Pinheiro, L. B. et al. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal. Chem.* **84**, 1003–1011 (2012).
19. McCarty, D., Monahan, P. & Samulski, R. Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene. Ther.* **8**, 1248–1254 (2001).
20. Zhong, L. et al. Self-complementary Adeno-associated Virus 2 (AAV)–T cell protein tyrosine phosphatase vectors as helper viruses to improve transduction efficiency of conventional single-stranded AAV vectors in vitro and in vivo. *Mol. Ther.* **10**, 950–957 (2004).
21. McCarty, D. M. Self-complementary AAV vectors; advances and applications. *Mol. Ther.* https://doi.org/10.1038/mt.2008.171 (2008).
22. Eisenhut, P. et al. Orthogonal characterization of rAAV9 reveals unexpected transgene heterogeneity. *J. Biotechnol.* **393**, 128–139 (2024).

Acknowledgements

We thank Betsy Derwinski from Lonza Morristown, Legal Counsel team for thorough reading of the manuscript and providing valuable input. We thank Jacob Mardick, a former member of Lonza Houston R&D team, for his technical suggestion and support. This work was performed at the viral vector R&D laboratory in Lonza Houston, Inc.'s facilities in Houston, Texas, USA and at Lonza Group Ltd.'s facilities in Basel, Switzerland.

Author contributions

Conceptualization, T.D., M.F., P.W., B.G.; Methodology, T.D., M.F, C.L, P.W., B.G.; Investigation, T.D., M.F., C.L; Writing, T.D., M.F, P.W.; All authors read and approved the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-77378-7.

Correspondence and requests for materials should be addressed to T.D., B.G. or P.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.