

# A dual-reference study design for understanding and improving AAV genome size analysis using ssDNA 7K ladder.

---

## Authors

Zhi-xiang Lu, Yali Sun, Michael Miller, Anis Khimani, Jana Bauer, Michael Salomon, Yanhong Tong

Revvity, Inc.

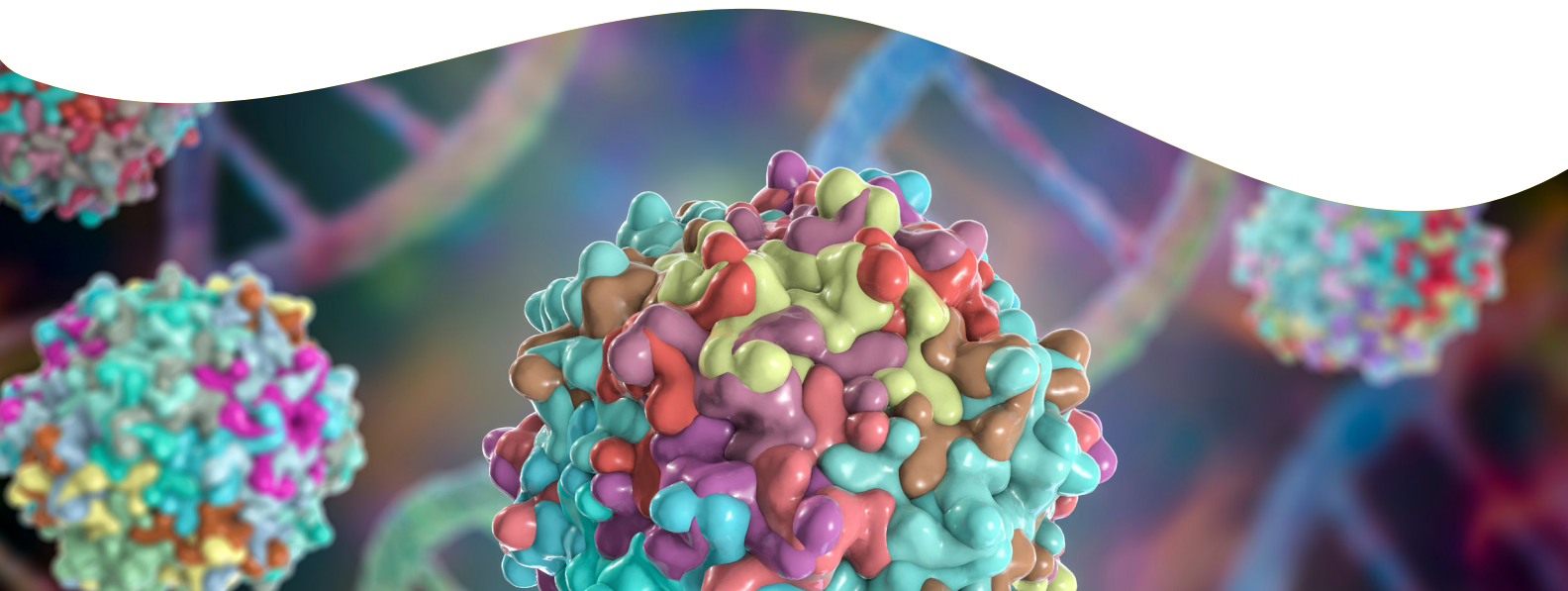
For research use only. Not for use in diagnostic procedures.

## Abstract

We recently developed a novel, large size single-stranded DNA (ssDNA) size standards ranging ~1k to 7.2k (called ssDNA 7K Ladder). This product is an important size standard for large-size ssDNA analytical method development. Here, we further employed this novel standard to investigate AAV genome migration in electrophoresis. We identified that AAV genome features and sample processing conditions can influence the migration patterns. We report our findings here to raise awareness of this observation from the current study. If unrecognized, researchers might mischaracterize the identity of DNA species in the AAV sample. Finally, we recommend a comprehensive design for AAV genome size integrity analysis using the ssDNA 7K Ladder.

## Introduction

Adeno-associated virus (AAV) is a non-enveloped virus that can be engineered to deliver DNA to target cells. The ability to generate recombinant AAV particles lacking any viral genes and containing DNA sequences of interest for various therapeutic applications has thus far demonstrated to be safe and effective, with four FDA approved gene therapies until mid-2023 (Luxturna 2017, Zolgensma 2019, Hemgenix 2022, and Elevidys 2023). Despite the documented AAV-therapeutic success, many preclinical and clinical studies have been still struggling with AAV delivery and tropism, transfection/transduction efficiency, long-term efficacy, and immunogenicity. The integrity and purity



of cargo DNA in AAV packaging affects clinical potency and safety, which is one of the most common challenges for AAV development and manufacturing quality control.

The wild-type AAV genome is a single-stranded DNA (ssDNA), either positive (plus) or negative (minus) sensed, with palindromic inverted terminal repeat (ITR) sequences forming double-stranded DNA (dsDNA) hairpin structures at each end. Based on vector engineering, there are two types of recombinant AAV vectors (rAAV): single-stranded AAV (ssAAV) and self-complementary AAV (scAAV). For ssAAV, the coding sequence and complementary sequence of the transgene expression cassette are on separate strands and are packaged in separate viral capsids. For scAAV, both the coding and complementary sequence of the transgene expression cassette are present on the same genome either plus or minus strand. The configuration of ssDNA and dsDNA structures emerge as a result of two rAAV genomes mentioned above. The ssAAV genome includes two ITRs (dsDNA region) flanking a transgene cassette (ssDNA region), whereas the scAAV genome includes the two ITRs flanking a transgene cassette (dsDNA region) and the third middle ITR dividing two strands of a transgene. In the current study described below, unless specified, all AAV samples used were ssAAVs.

Electrophoresis including gel electrophoresis, capillary electrophoresis (CE) and microfluidic CE (MCE) are simple and widely used methods to study the size of nucleic acids (DNA and RNA). The mobility of ssDNA or dsDNA in electrophoresis depends not only on its mass and shape but also on the electrophoretic conditions, such as gel/capillary components, running buffers, denaturing agents used, etc. Compared to scAAV which forms a dsDNA structure within the encapsulated virus particle, ssAAV presents multiple DNA conformations. Firstly, due to the ITR, the ssDNA is not a linear, open-ended molecule but rather forms T-shaped hairpin structures on both ends. Secondly, ssAAV packages either (+) or (-) strand with equal frequency. Following lysis (heat or proteinase K treatment), DNA is released from the virus particles, and subsequently dsDNA is formed via the annealing of the (+) and (-) strands during the sample preparation process for electrophoresis study. It is, therefore, important to consider factors such as the electrophoresis platform, DNA conformational and structural complexity as well as sample processing conditions when investigating the genome size of ssAAV by electrophoresis.

Here, we report a comprehensive study of the ssAAV genome migration patterns which can be an important reference to investigators in the gene therapy vector discovery and development area which is evolving rapidly. We have characterized the ssAAV ssDNA under key parameters utilizing the MCE platform, LabChip® GXII Touch™ protein characterization system. In addition, we have also used the ssDNA 7K ladder as a molecular size reference to evaluate the integrity of the ssAAV genome.

## Materials and methods

### INSTRUMENTS, REAGENTS, ssDNA 7K PROGRAM

- LabChip® GXII Touch™ protein characterization system (P/N CLS138160) and DNA 5K/RNA/CZE HT LabChip (P/N 760435)
- Agilent 2100 BioAnalyzer (P/N G2939BA) and DNA 12K kit (P/N 5067-1508)
- ssDNA 7K Ladder (P/N CLS157950)<sup>1,2,3</sup> and RNA Pico Assay Reagent Kit (P/N CLS960012)
- ssDNA 7K Assay Run Program (ssDNA 7K.asyx, P/N CLS158169)<sup>3</sup>
- Zymo Research Oligo Clean & concentrator Kit (P/N D4060)

### SAMPLES

- AAV premade eGFP control viruses (named as AAVx-GFP, full ssAAV genome size = ~3.3kb) and the AAV construct (named as pAAV-eGFP) were provided by Revvity SIRION-Biotech<sup>4</sup>.
- AAV reference standard was purchased from ATCC (AAV8-RSS, cat # VR-1816).

### dsDNA AND ssDNA SIZE STUDIES IN TWO MCE PLATFORMS

- dsDNA detection platform (labeled as “BioA dsDNA 12K assay” in figures): using Agilent 2100 BioAnalyzer with DNA 12K kit, samples without heat denature treatment, following the user guide from the vendor.
- ssDNA detection platform (labeled as “LabChip® ssDNA 7K assay” in figures): using LabChip® GXII Touch™ instrument together with ssDNA 7K Ladder and RNA Pico reagent, following the user guide from the vendor<sup>3</sup>.

## AAV SAMPLE PREPARATION: DENATURING AND ANNEALING PROTOCOLS

- Denature (Method-1): AAV samples were DNase I treated, lysed by Proteinase K using the Zymo Kit. The purified nucleic acid samples were denatured by heat treatment, 95°C for 3min, followed by snap cooling on ice for at least 5 min.
- Annealing (Method-2): DNA samples were denatured by heat treatment, 95°C for 10 min, subsequently slow cooling to room temperature for 20 min, followed by purification using the Zymo kit.

Table 1: Samples used in the study

Type	Labeling	Structure*	Expected size	Used in
Virus	AAVx-eGFP, or AAV-eGFP denature	ssDNA	3.3 kb	Figure 1, 2
Plasmid	pAAV-eGFP	dsDNA	5.9 kb	Figure 1, 2
DNA fragments	dsDNA-frag	dsDNA	2.1 kb, 5.1 kb	Figure 2
DNA fragments	dsDNA-frag denature	ssDNA	2.1 kb, 5.1 kb	Figure 2
DNA fragments	dsDNA-AAV-Ctrl	dsDNA	2.6 kb, 3.3 kb (full AAV genome with two ITRs)	Figure 2, 3
DNA fragments	dsDNA-AAV-Ctrl denature	ssDNA	2.6 kb, 3.3 kb (full AAV genome with two ITRs)	Figure 2
DNA fragment	dsDNA-AAV-Full	dsDNA	3.3 kb (full AAV genome with two ITRs)	Figure 2
DNA fragments	dsDNA-Ref	dsDNA	1.1 kb, 2.1 kb, 4 kb	Figure 4
Virus	ATCC-AAV8RSS	ssDNA	4.3 kb	Figure 4

\*In theory.

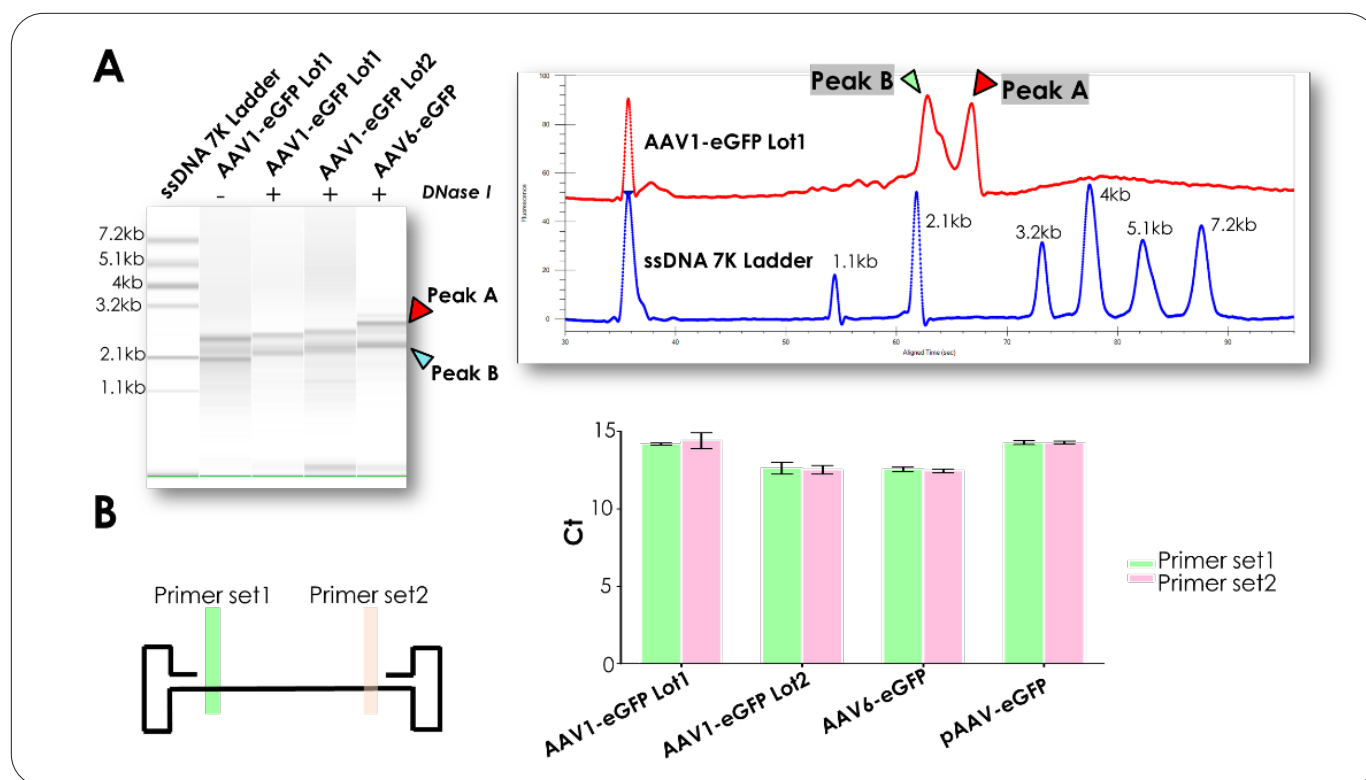
## Results

### Observation of two peaks in some AAV Genome Analysis

Our previous work has demonstrated the feasibility of AAV genome size evaluation using ssDNA 7K Ladder on LabChip® GXII Touch™ instrument<sup>2</sup>. The goal of this study is to refine our understanding and gain more insights into the migration of the AAV genome using the MCE platform. To assess the electrophoretic migration patterns, we began the evaluation with known materials: AAV control viruses packaged with the same GOI (Gene of Interest: eGFP, full AAV genome size=~3.3kb).

The control samples above were obtained from Revvity SIRION-Biotech which were purified and characterized by iodixanol gradient centrifugation, for the enrichment of full capsids, and subsequently titrated by qPCR using ITR

specific PCR primers.<sup>4</sup> Samples were lysed by proteinase K treatment, purified by Zymo kit. and heat denatured at 95°C for 3 min heat followed by snap cooling to 4°C before subjecting to MCE. Interestingly, two peaks were observed across all samples, irrespective of the lots of the viral preparations, or the viral serotypes, with or without pre-treatments such as DNase I (to eliminate host cell and residual plasmid DNA) (**Figure 1A**). Further analysis using qPCR showed Ct number of 5' end and 3'end of transgene's primer sets which were comparable (**Figure 1B**), indicating a two-peak migration pattern (corresponding to "Peak A" and "Peak B" as follows) is less likely related to partial ssAAV genome. Additional experiments were designed to characterize these two peaks.



**Figure 1: (A) Two-peak migration pattern of AAV control materials using ssDNA 7K Ladder on LabChip® GXII Touch™.** Left panel: Gel images of AAV reference samples of two different serotypes, lots, or DNase I treatments; Right panel: Electropherogram. All samples (1e13GC/mL virus, 50  $\mu$ L each) represented here with (+) or without (-) DNase I pretreatment were lysed by Proteinase K, purified by Zymo kit and eluted in 20  $\mu$ L TE (10mM Tris, pH=8). Two peaks observed and discussed in the text are labeled as Peak A and Peak B. **(B) AAV genome integrity analysis by qPCR using two primer sets.** Left panel: primer locations in the AAV genome; Right panel: qPCR results of AAV samples. pAAV-eGFP is AAV plasmid representing the full genome control. Error bars: standard deviations from duplicate tests.

## Characterization of two peaks in AAV Genome Integrity Analysis

### 1. Peak A is AAV full size ssDNA genome

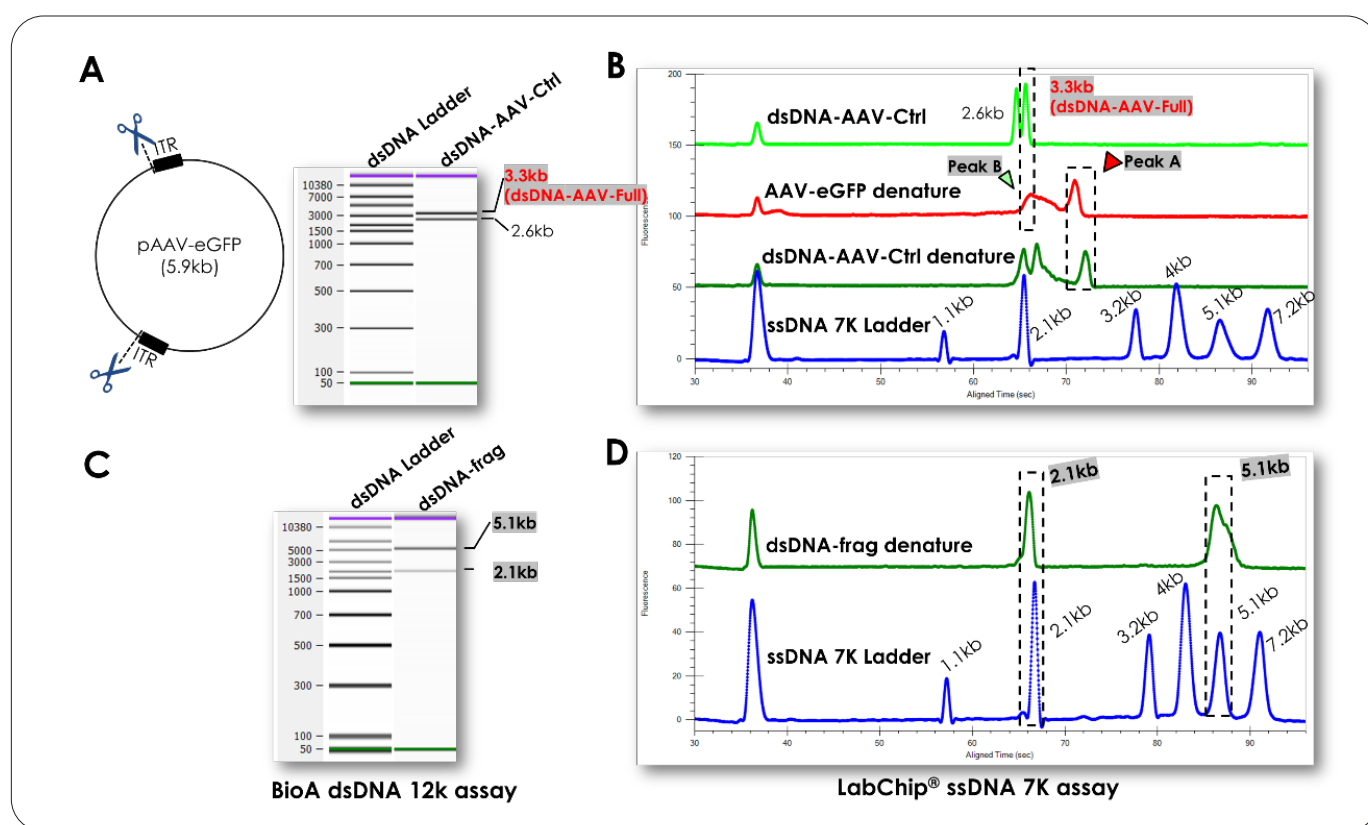
In our previous study, we had observed that AAV ssDNAs migrated faster (i.e., smaller size) compared to their expected sizes based on the ssDNA nucleotide lengths and corresponding detection in the MCE platform<sup>2</sup>. For instance, for the full capsid reference sample from Charles River Laboratories, expected size is ~2.4 kb, however the full genome peak is observed at ~2.1 kb ssDNA time point<sup>2</sup>. This observation led us to hypothesize that Peak A (observed size = ~2.9 kb) could also be the full size ssAAV genome (expected size = ~3.3 kb). To investigate this, we prepared one control nucleic acid sample (labeled “dsDNA-AAV-Ctrl”). Two restriction enzyme cleavage sites were carefully chosen just outside of ITR in the AAV plasmid (pAAV-EGFP) which were used for virus production (Figure 2A). This “dsDNA-AAV-Ctrl” has two dsDNA fragments, one is ~2.6 kb (49% GC), the other is ~3.3 kb (55% GC). As ~3.3

kb fragment contains the exact AAV full genome including two ITR regions (labeled as “dsDNA-AAV-Full”), it was treated as an ideal reference/comparison band or peak in electrophoresis analysis. The size of “dsDNA-AAV-Ctrl” was first confirmed by Agilent BioAnalyzer with DNA 12K kit, a dsDNA detection platform called “BioA dsDNA 12K assay” in Figure 2A. The AAV sample (red profile within the electropherogram in Figure 2B) and the “dsDNA-AAV-Ctrl” (dark green profile in Figure 2B, labeled as “dsDNA-AAV-Ctrl denature”) were treated with the same heat denaturation process; and subsequently analyzed by the LabChip® GXII Touch™ RNA Pico kit and using an independent size reference, the ssDNA 7K Ladder. As expected, Peak A, representing the AAV denatured sample migration behavior was similar to the reference/comparison band (denatured 3.3kb dsDNA-AAV-Full fragment) (Figure 2B, boxed peaks within the red and the dark green profiles). Peak A is the intact AAV ssDNA genome, evident from the faster-migrating pattern reproducing our findings of similar ssAAV genome analysis using the same MCE platform and the ssDNA

ladder as the size reference. We then asked whether ssDNA fragments (without ITR structures), other than AAV samples, appear smaller than their expected size when compared to the migration of the size reference, the ssDNA 7K Ladder. To investigate this, we prepared “dsDNA-frag” containing 2.1 kb and 5.1 kb dsDNA fragments (containing 44% GC and 42% GC content, respectively, without the ITR structures). The size of “dsDNA-frag” was confirmed by “BioA dsDNA 12K assay” (Figure 2C). After heat denaturation, ssDNA forms were detected, and corresponding sizes were determined by comparison to the migration pattern of the

ssDNA 7k ladder. As apparent in Figure 2D, the 2.1 kb, 5.1 kb denatured single-stranded bands were aligned well with 2.1 kb and 5.1 kb bands of the ssDNA 7K Ladder profile (highlighted by the dashed boxes). This result shows that the smaller observed size is ssAAV sample dependent.

In summary, the difference of observed size vs. expected size of AAV ssDNA genome could be related to the AAV sequence itself (GC content, secondary structure, etc.). Peak A (with relative slower migration, larger size to Peak B) is the intact AAV ssDNA genome.



**Figure 2: Peak A represents single-stranded, full AAV genome and its observed size is sequence dependent.** (A) 3.3 kb dsDNA-AAV-Full dsDNA-frag migration pattern on Agilent BioAnalyzer with dsDNA 12K kit. (B) Heat denatured 3.3 kb dsDNA-AAV-Full migration pattern on LabChip® GXII Touch™ instrument with RNA Pico Kit and ssDNA 7K Ladder. (C) 5.1 kb and 2.1 kb dsDNA-frag migration patterns on Agilent BioAnalyzer with dsDNA 12K kit. (D) Heated 5.1 kb and 2.1 kb dsDNA-frag migration patterns on LabChip® GXII Touch™ instrument with RNA Pico Kit and ssDNA 7K Ladder. Dashed boxes indicate certain peak migration comparisons between ssDNA ladder or control samples. For instances, in (B), Peak A observed in AAV sample migrates similarly compared to 3.3 kb dsDNA-AAV-Full denatured ssDNA peak (dark green). Peak B observed in AAV sample migrates similarly 3.3 kb dsDNA-AAV-Full non-denature dsDNA peak (Light green). dsDNA-AAV-Ctrl is a product of AAV plasmid pAAV-eGFP cleaved by restriction enzymes as shown in (A). ITR: inverted terminal repeats.



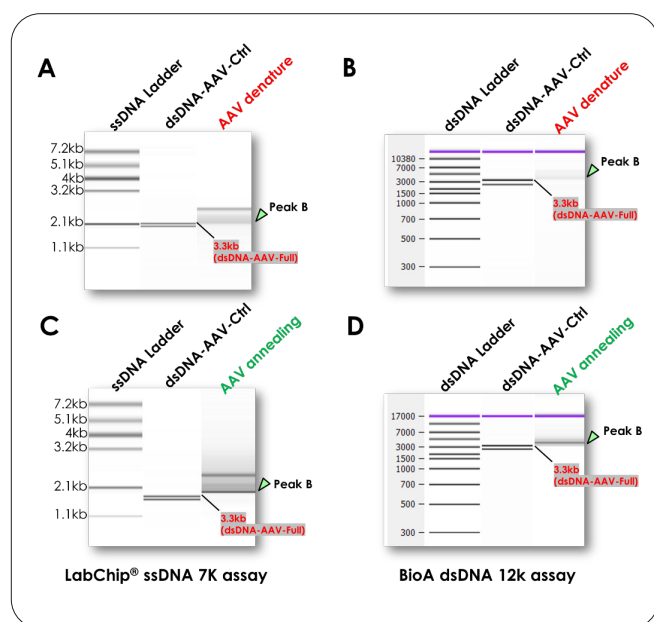
## 2. Peak B is AAV full reannealed dsDNA genome

Just as Peak A was aligned with the control 3.3kb “dsDNA-AAV-Full” denatured ssDNA peak, we observed a similar phenomenon with Peak B which migrated similar to the control non-denatured 3.3 kb “dsDNA-AAV-Full” dsDNA peak (**Figure 2B**, boxed red vs light green peak), indicating that Peak B is likely an annealed dsDNA structure from the plus strand and minus strand of AAV ssDNA. To test this assumption, we analyzed the “dsDNA-AAV-Ctrl” and AAV sample in both dsDNA detection MCE platform and ssDNA detection MCE platform (**Figure 3**). The double-strand detection MCE platform can only detect dsDNA structures with dsDNA specific intercalating dye, while the single-strand detection MCE platform can detect both ssDNA and dsDNA structures with ssDNA intercalating dye based on our previous observations<sup>1,2</sup>. “dsDNA-AAV-Ctrl” was directly analyzed in both platforms without denaturation. AAV sample was prepared by two methods as described below. AAV sample was prepared by denaturing protocol (Method-1 above, results in **Figure 3A-B**): AAV sample was DNase I treated to remove residual host DNA, lysed by Proteinase K then purified by Zymo Kit to remove protein and salts. The resulting purified nucleic acid sample was denatured by heat treatment (95°C for 3 min, followed by snap cooling on ice for at least 5 min) before loading on to the dsDNA detection platform (**Figure 3A**) and the ssDNA

detection platform (**Figure 3B**). AAV sample was prepared by annealing protocol (Method-2 above, results shown in **Figure 3C-D**): Lysed at 95°C for 10min, followed by slow cooling at room temperature for 20 min to facilitate annealing followed by purification using the Zymo kit to remove protein and salts. The same purified nucleic acid sample were directly loaded on to the LabChip® platform alongside the ssDNA 7k ladder (**Figure 3C**) and on to the Bioanalyzer 12k dsDNA assay platform (**Figure 3D**).

If Peak B is a dsDNA form, we would expect to see Peak B in the dsDNA detection platform. In **Figure 3B**, a lighter but noticeable band was detected at the expected location with an observed size of around 3.3 kb. Furthermore, if the denatured DNA representing Peak B annealed as a result of method-2 above to AAV dsDNA, we would expect to see an increased band intensity of Peak B. As apparent, upon annealing, Peak B was the predominant band as size referenced to the ssDNA 7k ladder evaluated via the LabChip® platform (**Figure 3C**). Furthermore, similar dense band was observed when evaluated by the Bioanalyzer 12k dsDNA assay (**Figure 3D**).

In summary, we have shown above that the faster migrating Peak B (relative to Peak A) is the intact AAV dsDNA genome per our investigation via the annealing process of plus and minus strands of AAV ssDNA.



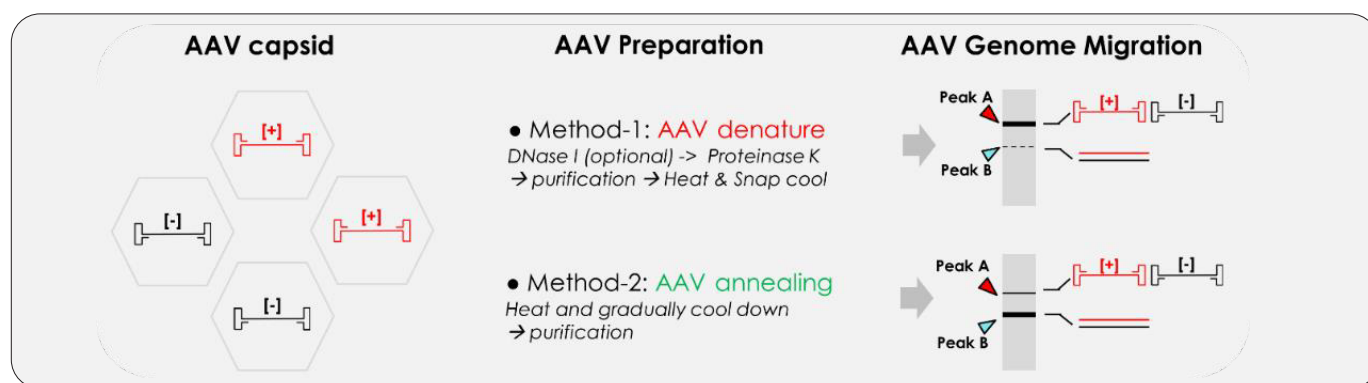
**Figure 3: Peak B represents the annealed double stranded AAV genome from AAV plus strand and minus strand. (A)(C) ssDNA detected using the LabChip® platform: Samples were size referenced using the ssDNA 7K Ladder on LabChip® GXII Touch™. (B)(D) dsDNA detection using Agilent Bioanalyzer platform: Samples were sized using Agilent dsDNA 12K kit. dsDNA-AAV-Ctrl is AAV plasmid cleavage product (3.3 kb + 2.6 kb) without heat denaturation as shown in Figure 2A. **AAV denature** (red legend): AAV sample was treated as described in the text. **AAV annealing** (green legend): AAV sample was treated as described in the text. Please note the same amount of start materials was used (1e13GC/mL virus, 50 µL each) in **AAV denature** and **AAV annealing**; After purification, both **AAV denature** and **AAV annealing** samples were eluted in 40 µL TE (10 mM Tris, pH=8).**

## Two-peak model under two AAV sample processing methods

Based on the results from the above study, we propose a two-peak migration model for the AAV genome size analysis using the LabChip® platform and the ssDNA 7K Ladder (**Figure 4**). AAV capsids are theoretically packaged either plus (+) or minus (-) strand of the DNA genome with equal frequency. AAV sample preparation process will affect the predominance of ssDNA or dsDNA genome structure formation. ssDNA and dsDNA genome structures migrate differently upon electrophoresis, appearing as two peaks as labeled Peak A (ssDNA form) and Peak B (dsDNA form), respectively. As shown above, in AAV preparation method-1, when samples are heat denatured and snap

cooled on the ice, Peak A (ssDNA form) is the only or major band. However, if researchers use method-2 annealing protocol as described above, double-stranded form is dominant, whereas the single-stranded form is minor, as depicted in **Figure 4**.

Please note that both method-1 and method-2 are standard AAV nucleic acid preparations. Depending on the sensitivity and sizing resolution of the electrophoresis platforms, it is normal to see either two-peak ("Peak A" and "Peak B") or single-peak (only "Peak A" or only "Peak B") pattern in ssAAV genome analysis. "Peak B" may be misinterpreted as a partial sized ssDNA genome without proper analysis or confirmation with orthogonal methods.



**Figure 4:** AAV genome migration patterns in electrophoresis by two AAV preparation methods.

## A framework of comprehensive design for AAV genome integrity analysis

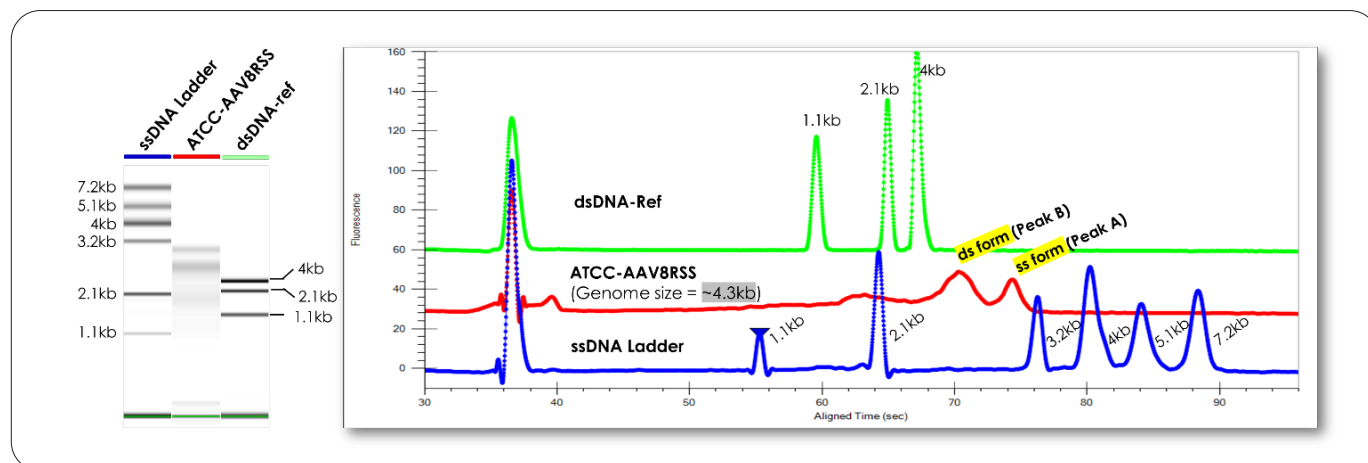
Considering an AAV genome in its configuration containing the T-shaped hairpin structures on both ends of the ssDNA, (-) and (+) strand genome annealing, AAV sample preparation, and electrophoresis conditions will greatly influence the final population of DNA species, resulting in complex migration patterns in electrophoresis analysis. Learnings from the above study, we suggest a framework of comprehensive design for AAV genome size integrity analysis on the MCE platform. In addition to the AAV nucleic acid purification procedure in the workflow as mentioned previously<sup>2</sup>, a known size dsDNA reference without heat denaturation is strongly recommended to be included as an additional sample along with the ssDNA 7K Ladder when using the MCE platform. We recommend that this dsDNA reference be a full AAV genome dsDNA

cleaved the same AAV construct used for AAV production as shown in **Figures 2-3**. However, other dsDNA fragments with a serial size representation from 1.1kb, 2.1kb and 4kb (labeled as "dsDNA-ref" in **Figure 5**) which covers typical AAV packaging size range can also be used as a dsDNA reference. This dsDNA reference will further enable size evaluation of annealed dsDNA structures from AAV sample, particularly from ssAAV.

Following this new design, we analyzed the ATCC AAV8 reference standard using method-1, representing AAV denaturation. As shown in **Figure 5**, two peaks were observed, a single-stranded form ("ss form": Peak A) at ~ 3 kb that migrates faster (representing shorter size) than its expected size of 4.3 kb, whereas the double-stranded form ("ds form": Peak B) correlates with ~4.3 kb dsDNA size expectation and migrates slightly slower than 4 kb dsDNA reference peak (ladder shown in green).

In summary, our dual-reference design (ssDNA ladder and dsDNA ladders) recommendation can enable characterization of intact ssAAV genome either in the ssDNA form or the dsDNA form. Bands that may appear in

other sizes besides Peak A and Peak B, especially smaller size fragments, will be considered impurities or partially packaged AAV genome.



**Figure 5: Representation of a comprehensive design for AAV genome integrity analysis.** Left panel: Gel images. Right panel: Electropherogram. ATCC - AAV8RSS samples were processed by method-1 without DNase I treatment as in Figure 4. Both ssDNA 7K Ladder and dsDNA size references (1.1 kb, 2.1kb, and 4kb) were used as single-stranded and double-stranded DNA sizing references.

## Conclusion and discussion

The study here addresses AAV genome characterization using microfluidic capillary electrophoresis (MCE), with specific focus on the LabChip® platform. During our research efforts, we report here a novel observation of twin or dual peaks. This appearance of the two peaks was consistent at the same size location across the electrophoretic profile when run and compared alongside two references, the ssDNA 7K Ladder (reported in previous studies<sup>1,2</sup>) and the dsDNA fragments. Based on this novel finding, we further investigated and characterized the two peaks which enabled us to understand AAV genome migration patterns. From the above study, we introduce a new “dual-reference” (size standards) concept that can be applied to multiple vector or carrier genome identification and characterization work across other electrophoretic technology platforms.

Recombinant AAV vectors have been widely utilized for *in vivo* gene therapy. Genetic modification of AAV vectors is further facilitating the success of AAV- vector-based gene therapy. Through multiple approached, such as AAV capsid

vector engineering increases delivery specificity (tissue tropism) as well as AAV transgene vector engineering can increase transduction efficiency and cargo expression. Srivastava et al. reported a 20-nucleotide sequence termed D sequence in ITR region impaired rescue, replication, and encapsulation of the viral genome. Their AAV vector engineering studies indicated the removal of D sequence, and replacement with a sequence containing putative binding sites for transcription factors in single-stranded AAV (ssAAV) vectors producing single-polarity AAV particle containing either (-) or (+) polarity genomes, and significantly augmenting transgene expression<sup>5</sup>. Our “dual-reference” ssDNA 7K Ladder and dsDNA size reference analysis method is capable of detecting AAV single polarity, full or partial sized AAV genome as well as small size impurities. Gaining an understanding through this study and other studies will enhance the analytical endeavor from early discovery (such as AAV vector engineering) through late stage CQAs (Critical Quality Attributes) across the entire workflow leading into manufacturing and QA/QC.



## References

1. Lu et al. 2023 Bioanalytical evaluation of single-stranded DNA for gene therapy applications
2. Lu et al. 2023 Considerations in AAV genome integrity analysis
3. ssDNA 7K Ladder
4. SIRION Biotech AAV Manufacturing Service
5. Ling et al. 2015 Enhanced Transgene Expression from Recombinant Single-Stranded D-Sequence-Substituted Adeno-Associated Virus Vectors in Human Cell Lines In Vitro and in Murine Hepatocytes In Vivo J Virol. 89(2):952-961

## Technical contact information

Please see instrument, reagent and software section listed above for ordering. Please send questions or requests for user guide and/or evaluation reagent to: [dxsupportamericas@perkinelmer.com](mailto:dxsupportamericas@perkinelmer.com) by adding "ssDNA-" at the beginning of the subject line.

The Revvity logo is displayed in a lowercase, sans-serif font. It is positioned on the right side of the page, above a large yellow wavy graphic that spans the bottom of the document. The logo itself is black.