Quality Control for Adeno-Associated Virus Genome Preparations Using Nanopore Sequencing

Download poster

Hayley Greenfield[1], Heather Jeffery[1], Sissel Juul[2], Phillip James[1]

[1]Oxford Nanopore Technologies plc, Oxford, UK.
[2]Oxford Nanopore Technologies Inc, New York City, NY.

More information at: www.nanoporetech.com and publications.nanoporetech.com

Oxford NANOPORE Technologies™ APPLICATIONS

Abstract

Adeno-associated viruses (AAV) vectors are a gene delivery vehicle used in gene therapy. They are replication-deficient viruses with a 4.8 kb single-stranded DNA genome flanked by inverted terminal repeats (ITRs). AAVs are increasingly being used for *in vivo* DNA delivery, with five AAV-based gene therapies already approved by the FDA[1]. Identifying vector genome quality issues that can arise during the manufacturing process, such as mutations, truncation hotspots, and potential contaminants, is critical.

Here we describe an AAV workflow, developed by Oxford Nanopore Technologies, for Quality Control (QC) and validation of rAAV preparations and the plasmids used to create them. Nanopore sequencing can generate reads that span the entire AAV genome, from ITR to ITR, with the ability to sequence through GC-rich regions, capture vector heterogeneity, and identify nucleic-acid impurities from plasmid and cell line DNA. This allows for the full characterisation of the composition of rAAV vector preparations.

Background

The production of rAAV involves first making the vector containing the transgene of interest, flanked between two ITRs that are necessary for replication and packaging of the vector[2]. Replacement of the rep-cap genes, which were necessary for capsid production and replication, with a transgene of interest can produce two recombinant AAV (rAAV) vectors; single stranded AAV (ssAAV) and self-complementary AAV (scAAV), dependent on the size of the inserted gene(s) (Fig. 1a). The vector is then co-transfected with rep-cap and helper plasmids into cell lines for AAV production[3] (Fig. 1b). Within these production stages, mutations in the transgene, truncated AAV genomes and contamination of host cell or plasmid DNA can affect the efficacy of the final rAAV preparations[4]. Quality Control (QC) steps can be performed using nanopore sequencing on both the plasmids (Fig. 1c) and rAAV preparations (Fig. 1d) to assess the integrity at both of these stages in the process.

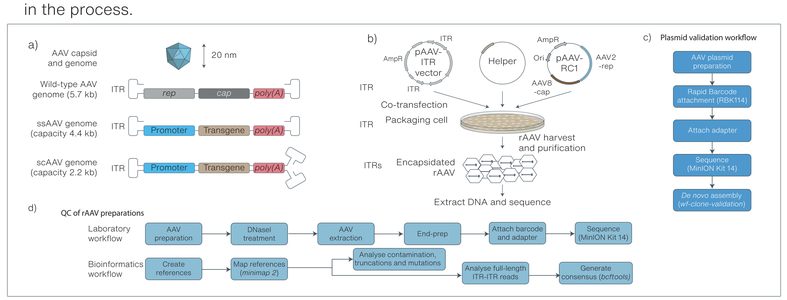


Fig. 1 AAV a) structure b) gene-therapy workflow c) plasmid validation workflow d) library prep and bioinformatic workflow for rAAV sequencing

Initial method optimisations revealed that there were significantly more full length AAV genomes sequenced when a direct ligation approach was taken. This was compared to both high (95°C) and low (70°C) temperature annealing protocols, with both annealing approaches resulting in a reduction of full length scAAV and ssAAV genomes (Fig. 2a-c).

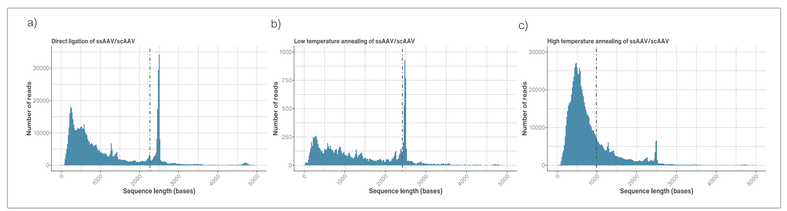


Fig. 2 Read length distributions for AAV8 direct ligation or annealing a) direct ligation b) low temperature annealing c) high temperature annealing

Plasmid sequencing

Plasmid sequencing was performed as an initial QC check and to provide the expected truth of the rAAV sequence. The earlier in the process deviations from the expected sequence are identified, the more cost-efficient the process can be.

Using the EPI2ME *wf-clone-validation* workflow and associated protocol, we sequenced a typical control AAV transgene plasmid containing the GFP gene (pAAV-GFP) using the Rapid Barcoding Kit (SQK-RBK114) on a MinION Flow Cell (R10.4.1). Read length distributions show a peak at ~ 5.3 kb indicating a single-cut full-length pAAV-GFP plasmid (Fig. 3a).

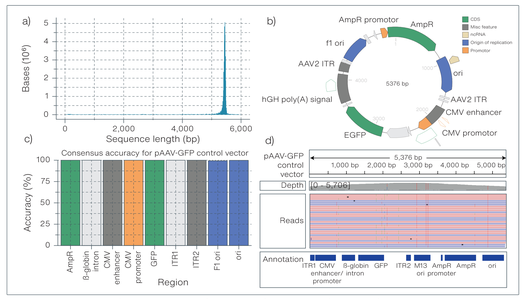


Fig. 3 Plasmid sequence a) read length b) plasmid map c) accuracy d) coverage

We used the EPI2ME *wf-clone-validation* workflow to create a plasmid map identifying the presence of the transgene and ITR sequences (Fig. 3b). This also identified all key plasmid components with high accuracy when compared with the manufacturer's reference (Fig. 3c). Alignments show even coverage across the plasmid indicating that circular plasmids were cut once with the rapid barcoding transposase (Fig. 3d).

Nanopore sequencing acts as a quality control tool for plasmids and AAV enabling the detection of contamination and truncations in multiplexed samples.

Sequencing of six multiplexed AAV samples

Multiplexing facilitates a higher number of samples to be sequenced in parallel. Due to the size of the AAV genome, multiplexing can provide a resource-efficient and cost-effective approach to AAV QC.

Two replicates of three serotypes, AAV5, AAV6 and AAV8, were extracted using the Purelink™ Viral DNA/RNA Mini Kit, barcoded and sequenced using the Native Barcoding Kit (SQK-NBD114) on a single MinION Flow Cell (R10.4.1). Read distributions show clear peaks for single-stranded AAV (ssAAV) and self-complementary AAV (scAAV) genomes as a mixed population, plus truncated genomes and potential contamination (Fig. 4a). Within one hour of sequencing, we obtained more than 200x coverage of the ITR-ITR region for each barcoded sample (> 95% read and reference coverage, Fig. 4b). Coverage of the rAAV vector was even across the ITR-ITR region for each barcoded sample (Fig. 4c).

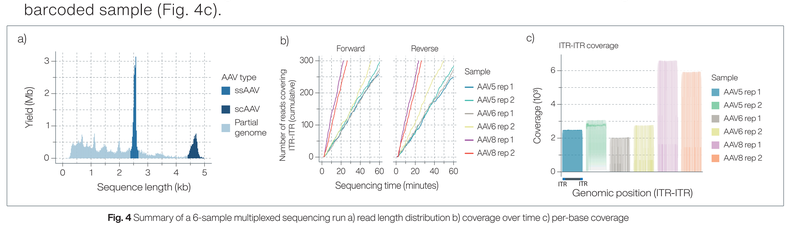
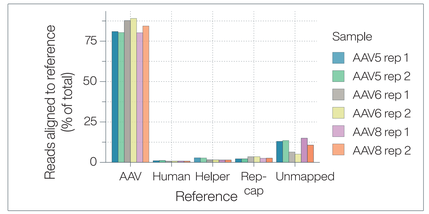


Fig. 4 Summary of a 6-sample multiplexed sequencing run a) read length distribution b) coverage over time c) per-base coverage

Detection of rAAV vector contamination

Using provided references for the human genome as well as rep-cap and helper plasmids, we were able to identify potential sources of contamination in the AAV sample (Fig. 5). A large proportion correctly aligned to the AAV reference generated during the plasmid sequencing step (Fig. 3b).



However, contaminating sequences aligning to human and both helper and rep-cap plasmid references were observed. We also noted several reads that did not map to the references used. These incorrectly packaged impurities in the sample may affect the efficacy of the final vector preparation.

Fig. 5 Read alignments to given references identifying level and type of contamination

Detection of truncation hotspots in AAV preparations

Identifying truncation hotspots and the generation of a consensus sequence of the AAV genome are common requirements for QC of AAV preparations. Calculating the sum of the start and stop positions of primary read alignments can identify truncated genomes and their locations.

Here we show that the AAV5 sample has a truncation in the CMV promoter which is absent in the AAV8 sample (Fig. 6a). Accuracy of the bcftools-generated consensus sequence for the AAV8 genome showed a reduction in the first ITR sequence (Fig. 6b) when compared against the supplied reference. This was due to an incorrect ITR reference and upon correction accuracies increased (Fig. 6c), demonstrating the importance of performing sequencing-based QC of the plasmid prior to rAAV production.

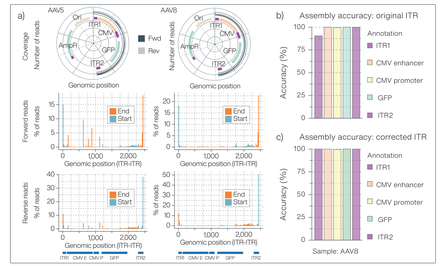


Fig. 6 a) coverage b) consensus accuracy before and c) after reference correction

Conclusion

AAV is increasingly being used as a favorable gene therapy vector for clinical applications, however contaminations in rAAV preparations can cause downstream adverse affects such as hepatotoxicities and neuropathological issues. Here we present a rapid, comprehensive workflow for the sequencing of rAAV vectors, allowing an accurate view of the composition of rAAV preparations contributing to the overall QC of these vectors. We have also demonstrated the option to multiplex AAV samples, providing a cost-efficient AAV QC method.

Workflow / protocol

References

[1] Au et al., 2022, "Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings". Frontiers in medicine, 8, 809118.
[2] Wörner et al., 2021, "Adeno-associated virus capsid assembly is divergent and stochastic". Nature Communications 12,1642.
[3] Xiao and Samulksi, 1998, "Production of high-titre recombinant adeno-associated virus vectors in the absence of helper adenovirus". Journal of Virology, 72(3), 2224-2232.
[4] Wang et al., 2019, "Adeno-associated virus vector as a platform for gene therapy delivery". Nature reviews drug discovery, 18(5), 358-378.