The Long and Short of It: Comparing Next Generation Sequencing Methods for Quality Control in AAV Development

Andrea O'Hara[1], Elizabeth Louie[1], Anusha Sriraman[2], Alpha Diallo[2], Eric Talevich[2], Amicia D Elliott[2], Haythem Latif[1], and Ginger Zhou[1]

[1]GENEWIZ from Azenta Life Sciences, South Plainfield, NJ 07080

[2]Form Bio, Austin, Texas 78701

Abstract

Background: Interest in gene therapy-based disease prevention and treatment has grown rapidly over the last decade, with adeno-associated viral vectors (AAV) emerging at the forefront with widespread application due to the non-integrating ability. With safety as the top priority in all therapies, extensive quality control (QC) throughout the entire development and manufacturing process is essential. A robust QC process expedites safe and effective commercialization of the final product. While Sanger sequencing can be ideal to verify and validate plasmid sequences pre-packaging, next generation sequencing (NGS) combined with post-viral production methodologies such as capillary electrophoresis (CE) analysis and transmission electron microscopy offer an effective high-throughput approach for monitoring AAV quality, from initial construct assembly to analysis of the encapsulated product.

Methodology: Illumina® short-read, Oxford Nanopore® long-read and PacBio® long-read sequencing technologies each offer distinct advantages including sequencing of the entire viral genome, with detection of potential mutations, truncations, and contaminants. Here, we compare PacBio Revio and Oxford Nanopore GridION as a means to assess the purity, clonality and fidelity of the packaged AAV product.

Results: When evaluating the same sample with Oxford Nanopore GridION and PacBio Revio, we show how these sequencing methods compare with detection of contaminating alternatively packaged products, detection of point mutations and indels in the ITRs and region of interest, and detection of full-length vs partially packaged genomes. Based off these results, we also evaluate how sample quality and library preparation can influence the conclusions generated and describe when each method can be ideally used within the therapeutic development workflow, starting from early research to late-stage pre-clinical work. Additionally, we consider regulatory assays, with Good Laboratory Practices (GLP) methods being ideal for late-stage development products.

Conclusion: While both methods prove to be comparable, a combined NGS approach post-packaging enables a comprehensive solution and enhances the overall QC process, ideal for sequence confirmation of both transfer plasmid and final packaged product for improved viral vector gene therapy manufacturing in advance of FDA or EMA filings.

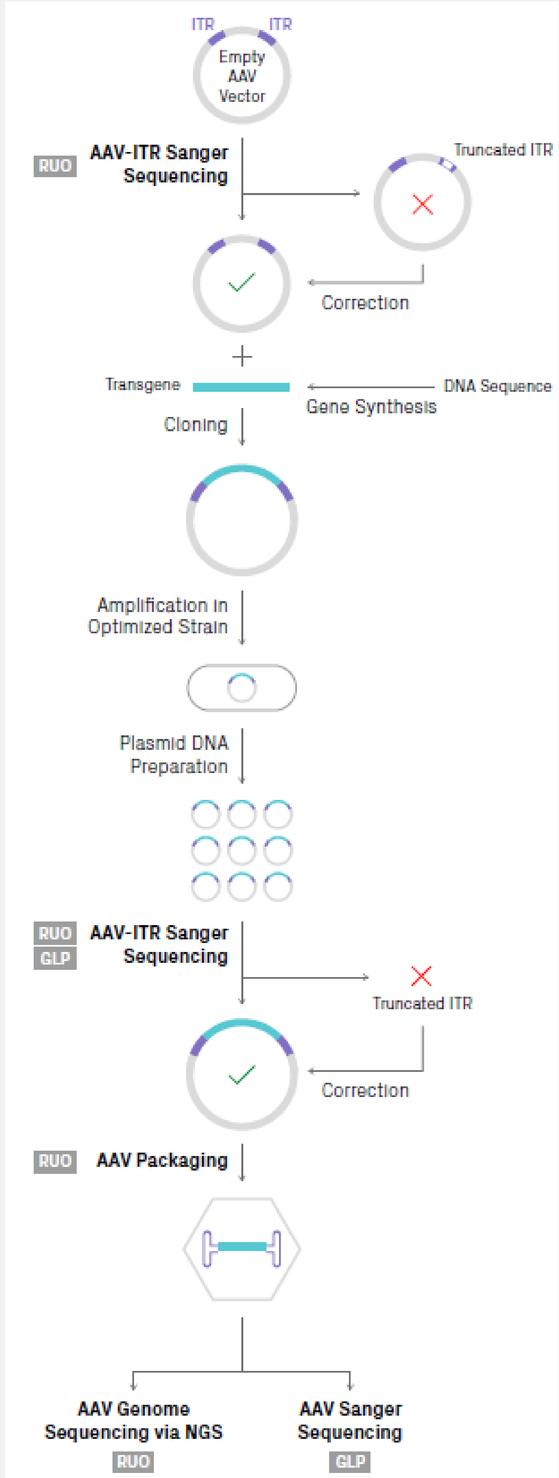
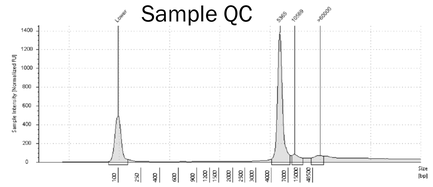


Figure 1. Full AAV synthesis and packaging workflow. Beginning with either full plasmid synthesis or transgene synthesis, the resulting transfer plasmid is amplified in an optimized strain to limit mutation and fully sequenced, including ITRs. Next, the transfer plasmid is used for AAV packaging. The resulting viral particles are subjected to further AAV genome sequencing with NGS and/or Sanger to confirm AAV sequence, verify AAV genome integrity, and identify and quantify any potential contaminants, including a partially packaged product and incorrectly packaged plasmid or host vector sequence.

AAV ssDNA is extracted from viral particles, along with any alternately packaged products (partial AAV genomes, residual host cell DNA, plasmid/vector DNA). Primary QC includes gel imaging to indicate general sample purity. Following QC, second strand conversion of ssDNA occurs, followed by independent library preparations for short- and long-read sequencing. Comprehensive post-sequencing analysis reports frequency of variants and contaminants along with size resolution of any contaminants and fidelity of final AAV genome including ITRs.

Oxford Nanopore GridION

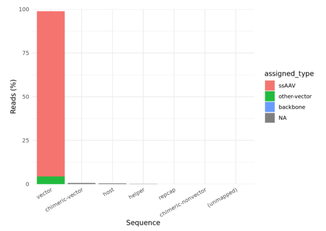
Sample QC



PacBio Revio

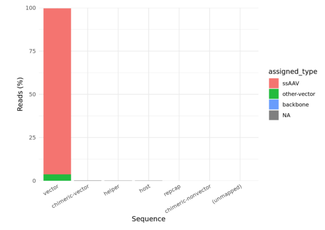
Assigned types by read alignment characteristics

Mapped Reference	Assigned Type	Count	Frequency (%)
vector	ssAAV	85,299	94.48
vector	other-vector	3,939	4.36
vector	backbone	12	0.01
chimeric-vector		581	0.64
host		326	0.36
helper		115	0.13
repcap		6	0.01
chimeric-nonvector		1	0.00
(unmapped)		4	0.00

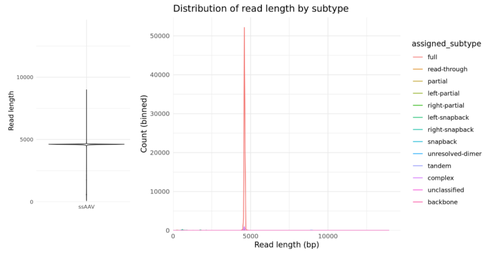


Assigned types by read alignment characteristics

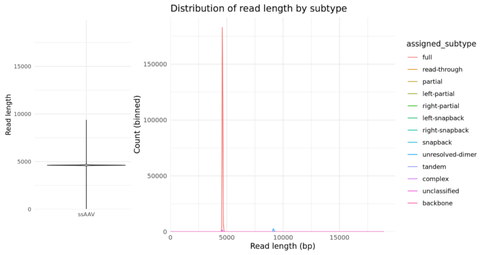
Mapped Reference	Assigned Type	Count	Frequency (%)
vector	ssAAV	324,890	96.00
vector	other-vector	12,380	3.66
vector	backbone	23	0.01
chimeric-vector		656	0.19
helper		272	0.08
host		191	0.06
repcap		4	0.00
chimeric-nonvector		3	0.00
(unmapped)		3	0.00



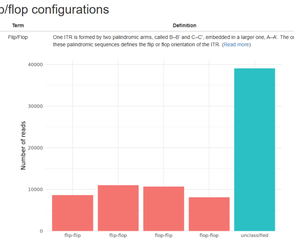
Distribution of read length by subtype



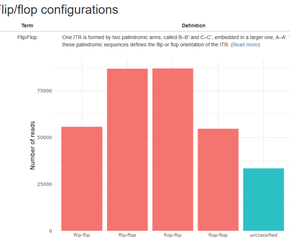
Distribution of read length by subtype



Flip/flop configurations



Flip/flop configurations



Preliminary QC indicates a highly pure packaged product, with no obvious evidence of alternatively packages products outside of the expected size. Post sequencing analysis with Form Bio's LAAVA pipeline from processing on Oxford Nanopore GridION (left) and PacBio Revio (right) present comparable results. Unsurprisingly, while the Revio sequencing is deeper, all trends remain in line regardless of sequencing platform. LAAVA analysis includes: assignment type by read alignment characteristics (including vector, chimera, RepCap, host and helper), and their read lengths, distribution of subtypes (including full, partial, and snapbacks) and their read lengths, and AAV mapping statistics to the reference sequence, including substitutions, insertions, deletions and ITR flip/flop configurations. Aside from read depth, the only notable difference between methods is the Flip/Flop configurations, which may be attributable to greater sampling with PacBio Revio.

Conclusions

Dual testing with Oxford Nanopore GridION and PacBio Revio yielded comparable results.

Advanced data analysis with Form Bio's LAVAA pipeline, yields comprehensive evaluation of packaged AAV products with fine resolution of sequence variants present within the AAV genome, along with the overall fidelity of the genome, including ITR regions.

This approach is ideal for testing of packaged product throughout the development and manufacturing life cycle.