Repeatability and Intermediate Precision of a Genome Integrity Assay for rAAV by Long-Read Sequencing

E. Lecomte[1], S. Maestro[1], A. Miguel Coello[1], O. Aldasoro Zabala[1], DL. Bowie[1], S. Stinus[1], I. Nechaev[1], N. Pawlowski[2], C. Trigueros Fernandez[1], A. François[1]
[1]Viralgen Vector Core S.L., San Sebastián, Gipuzkoa, 20009, Spain.
[2]Bayer AG, Analytical Development & Clinical QC, Wuppertal, 42117, Germany.

Introduction

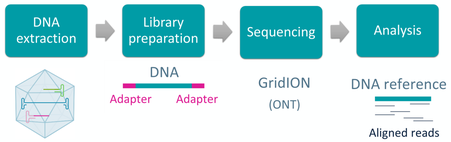
Truncated recombinant AAV genomes can compromise transgene expression and the therapeutic efficacy[1] of gene therapy vectors, making genome integrity a critical quality attribute, highlighting the need for robust analytical methods.

Long-read sequencing platforms, such as Oxford Nanopore Technologies (ONT), enable direct, full-length rAAV genome analysis and offer a powerful approach for genome integrity assessment[2,3]. Establishing key performance characteristics, like repeatability (intra-assay precision) and intermediate precision, becomes essential for use in quality control and release-testing.

This study assessed the repeatability and intermediate precision of a long-read sequencing-based genome integrity assay using replicate control DNA and self-complementary AAV (scAAV) samples across independent preparations.

Materials & Methods

Two samples were used to assess the precision of the integrity method: (i) a control DNA mixture composed of 92% full-length DNA, 5% truncated DNA, and 3% deleted DNA prepared using plasmid DNA restriction fragments, and (ii) a self-complementary rAAV (scAAV-GFP). Six independent sequencing runs were performed on the GridION platform (Oxford Nanopore Technologies), on different days, by different analysts, and using different flow cells. One replicate was sequenced per run for the control DNA mix, and one to six replicates were sequenced for the scAAV sample.



Assay (A)	Control DNA mix replicate (C)	scAAV replicate (S)	Day	Analyst	ONT Flow Cell
A1	A1-C1	A1-S1, A1-S2, A1-S3	1	1	1
A2	A2-C1	A2-S1	2	2	2
A3	A3-C1	A3-S1	3	2	3
A4	A4-C1	A4-S1, A4-S2, A4-S3, A4-S4, A4-S5, A4-S6	4	2	4
A5	A5-C1	A5-S1	5	3	5
A6	A6-C1	A6-S1	6	3	6

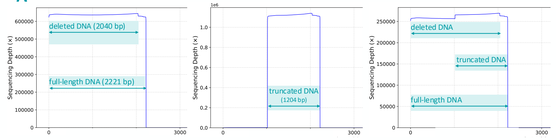
Precision evaluation was based on read-population profiles defined by alignment start and end positions. Populations were normalized by excluding low-abundance fragments (<0.5%) and focusing on hotspot alignments (≥3%).

Mean, standard deviation (SD), coefficient of variation (CV), and confidence intervals (CI) were calculated for both sample types using compositional-data methods based on log-ratio transformations.

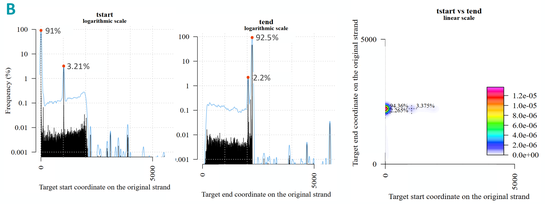
Control DNA Integrity Precision

To evaluate the precision of the integrity assay, a 2,221 bp DNA fragment was used as the full-length control DNA, and a truncated form of 1,204 bp was generated by restriction enzyme digestion and spiked into the full-length control at 5%. The original control DNA includes a 3% deleted form of 2,040 bp, which is detected by the truncation analysis as a secondary population appearing near the end of the sequence.

A



B



C

DNA sequence length	Characteristics consistent with	A1-C3	A2-C3	A3-C3	A4-C3	A5-C3	A6-C3
2221	Full-length DNA	94.4	92.7	92.4	92.5	91.9	93.7
1204	Truncated DNA	3.4	5.1	5.4	5.3	5.9	4.1
2040	Deleted DNA	2.3	2.2	2.2	2.2	2.2	2.2

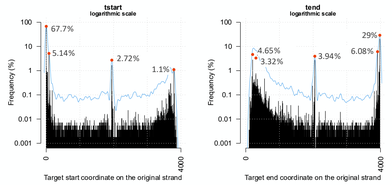
Fig.1 | Detection and Quantification of Truncation Events in Control DNA

(A) ONT sequencing depth across the DNA molecules for full-length (left), truncated (middle), and the mixed control A1-C1, consisting of 92% full-length DNA, 5% truncated DNA, and 3% deleted DNA. (B) Alignment frequency of read start sites (left), end sites (right), and distribution of the three read populations in the control A1-C1. (C). Relative quantification of read populations detected by the truncation analysis. Low-abundance fragments (<0.5%) were excluded, and only hotspot alignments were considered for normalization. A 2% hotspot threshold was applied to the control DNA to enable the relative quantification of the DNA containing the deletion. DNA populations were recovered at their expected proportions in the DNA control C1, in each experiment.

Expected start and end positions of full-length, deleted, and truncated DNA sequences are recovered in the control DNA. **The three populations are correctly resolved with expected proportions.**

rAAV Integrity

A



B

DNA sequence length	A4-S5
204	3.4
302	2.4
2051	2.5
3914	1.55
3981	**22**
3834	1.77
2034	1.8

C

DNA sequence length	A1-S1	A1-S2	A1-S3	A2-S1	A3-S1	A4-S1	A4-S2	A4-S3	A4-S4	A4-S5	A4-S6	A5-S1	A6-S1
204	3,29	4,33	3,52	4,07	4,83	3,7	3,47	3,16	3,03	3,37	3,41	3,37	4,29
268	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	4,33
298	ND	ND	ND	2,21	ND	ND	ND	ND	ND	ND	ND	ND	ND
302	1,86	2,11	1,89	ND	2,33	2,49	2,52	2,27	2,19	2,37	2,51	ND	ND
344	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	4,33	3,78
556	ND	ND	ND	0,802	ND	ND	ND	ND	ND	ND	ND	ND	ND
1134	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	0,75	ND
2051	1,84	1,6	1,91	0,98	0,832	2,01	2,25	2,48	2,69	2,52	2,6	2,64	2,25
3914	1,37	1,21	1,63	0,824	0,771	1,17	1,19	1,34	1,32	1,55	1,39	0,957	0,902
3981	**22,1**	**19,8**	**23**	**15,1**	**12,8**	**18,2**	**18,3**	**20,3**	**23,3**	**22**	**19,7**	**18,8**	**19**
124	ND	0,551	0,521	0,746	1,12	0,585	0,557	ND	ND	ND	ND	ND	ND
3834	2,61	2,21	2,78	1,38	1,93	1,7	1,65	1,52	1,92	1,77	1,37	ND	ND
2034	1,47	1,38	1,79	1,18	0,609	1,68	1,67	1,81	1,92	1,8	1,74	1,56	1,59
124	0,538	0,802	0,574	0,858	1,1	0,545	0,606	ND	0,504	ND	ND	ND	ND

Fig.2 | Detection and Quantification of Truncation Events in the rAAV Sample

(A) Alignment frequency of read start sites (left) and end sites (right) in the scAAV A4-S5 sample. (B) Read populations detected in the scAAV A4-S5 sample and (C) across all replicates. Low-abundance fragments (<0.5%) were excluded (ND). Only populations present ≥0.5% in every replicate (highlighted in green) were used to evaluate the intermediate precision. For repeatability (Assay 4), the 302-bp and 3834-bp populations (highlighted in orange) were included as they appeared in all replicates.

Read populations with characteristics common to distinct rAAV populations were identified:

1. **A population of 3981 bp, corresponding to the full-length scAAV**. The few missing bases at the beginning or the end of the alignments **(populations of 3834 bp and 3914 bp) are due to suboptimal alignments** related to the flip and flop ITR conformations present in the scAAV sample.
2. **Two populations of 204 bp and 302 bp, corresponding to truncated AAV sequences.**
3. **Two populations of 2051 bp and 2034 bp, corresponding to ssAAV (transgene monomers).**

Integrity Method Performance - Precision

A

Control DNA mix DNA characteristics consistent with	Mean (%)	SD (%)	CV (%)	95% CI of CV (%)
Full-length DNA	92.17	3.59	3.89	[1.3,7.4]
Truncated DNA	5.40	2.92	54.04	[21.1,85.4]
Deleted DNA	2.42	1.09	44.85	[17.6,71.6]

B

scAAV DNA characteristics consistent with	Repeatability – Assay 4				Intermediate precision			
	Mean (%)	95% CI of mean (%)	SD (%)	CV (%)	Mean (%)	95% CI of mean (%)	SD (%)	CV (%)
Truncated DNA, centered in the ITR sequence	10.18	[9.2,11.3]	1.3	12.81	13.72	[11.6,16.0]	4.09	29.82
ssAAV genomes	7.25	[6.6,7.9]	0.85	11.74	N.A.			
	5.33	[5.2, 5.5]	0.15	2.74	5.56	[4.9,6.3]	1.32	23.73
	7.29	[6.9,7.7]	0.49	6.69	7.37	[6.1,8.8]	2.49	33.77
Suboptimal alignmnent due to flip and flop conformations, and representing full-length scAAV	3.99	[3.8,4.2]	0.29	7.19	4.29	[3.9,4.7]	0.79	18.42
	4.99	[4.6,5.4]	0.52	10.37	N.A.			
Full-length scAAV	**60.97**	**[59.3,62.6]**	**2.06**	**3.38**	**69.07**	**[67.6,70.4]**	**2.57**	**3.73**

Fig.3 | Repeatability and Intermediate Precision Assessment

Mean, Confidence Interval of mean (CI), Standard Deviation (SD) and Coefficient of Variation (CV) used to evaluate the precision of the integrity method. (A) Control DNA mix results. (B) rAAV sample results.

Control DNA populations were correctly identified, with standard deviation ranging from approximately 1% to 4% (intermediate precision). CV rose with decreasing abundance, from 4% for the major control DNA population to 45–54% for minor species. **scAAV showed the same trend, with CVs of ~3–4% for the major population and ~18–34% for minor ones.**

Discussion & Conclusion

The integrity assay aggregates start and end positions of read alignments to identify and quantify full-length genomes and potential truncation hotspots (exact start and end positions not shown in the poster) across rAAV genomes. Control DNA populations were identified as expected, and distinct DNA populations were detected in the scAAV sample.

The integrity assay demonstrated consistent precision across runs, supporting its suitability for future validation.

1. McColl-Carboni, et al. (2024). Analytical characterization of full, intermediate, and empty AAV capsids. Gene Ther 31, 285–294.
2. Tai PWL, et al. (2018). Adeno-associated virus genome population sequencing achieves full vector genome resolution and reveals human-vector chimeras. Mol. Ther. Methods Clin. Dev. 9, 130–141.
3. Namkung S, et al. (2022). Direct ITR-to-ITR nanopore sequencing of AAV vector genomes. Hum. Gene Ther. 33(21–22), 1187–1196.

EUSKO JAURLARITZA GOBIERNO VASCO
Europar Batasunak kofinantzatua
Cofinanciado por la Unión Europea