# The Good, the Bad, & the Chimeric: Using Nanopore Sequencing to De-Risk rAAV Genome Packaging

K BREUNIG, F DUNKER-SEIDLER, M HAUBNER, F SONNTAG, M HOERER, RC FEINER

ascend
ABL inc.
ADVANCED BIOSCIENCE LABORATORIES

Poster 829

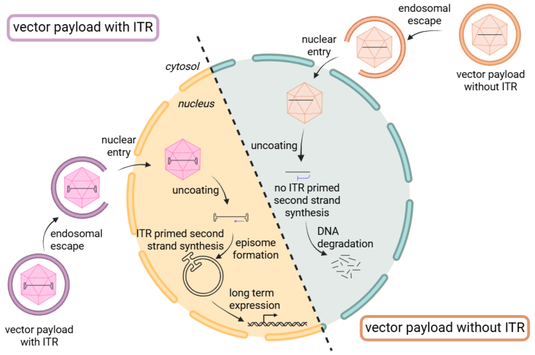
## Introduction

Recombinant adeno-associated virus (rAAV) is a widespread vector for gene therapy. The most common rAAV manufacturing process is based on transient transfection of HEK293 cells with plasmids carrying the vector genome cassette, the viral capsid (cap) and replication (rep) genes and adenoviral helper genes. During rAAV production, most capsids are equipped with the desired, full length vector genome ("the good"), however a small proportion of viral vectors contain different DNA payloads from the manufacturing plasmids and the host cell genome ("the bad"). These product-related impurities can trigger immune responses limiting therapeutic efficacy or even lead to severe adverse events linked to patient liver injury[1]. The potential to express DNA impurity encoded genes upon transduction is governed by the presence of inverted terminal repeats (ITRs). These sequences guide second-strand synthesis in the nucleus and episome formation enabling sustained gene expression.

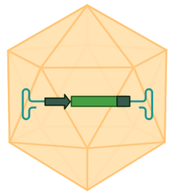
A priori, critical impurities such as helper genes or host cell DNA do not contain ITR sequences resulting in lacking or only transient expression. However, recombination events can lead to vector payloads containing ITRs fused to rep, helper or host cell DNA enabling long term expression with higher associated risk ("the chimeric"). The characterization of these events is challenging by common analytics such as alkaline gel or digital PCR. Here, we describe how short and long read sequencing can elucidate rare genomic fusion events and discuss their biological implications. As example data set, we used a research grade rAAV9-SEAP batch produced in a suspension HEK293 system combined with a single step, affinity-based purification method.

[1] Buddle et al. (2025), medrxiv, doi: 10.1101/2025.01.13.25320105

## ITRs enable long term transgene expression

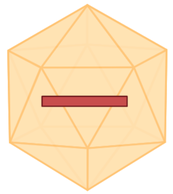



## AAV vector batches contain heterogenous sequences



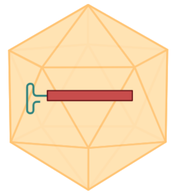
THE GOOD

The desired full-length vector genome



THE BAD

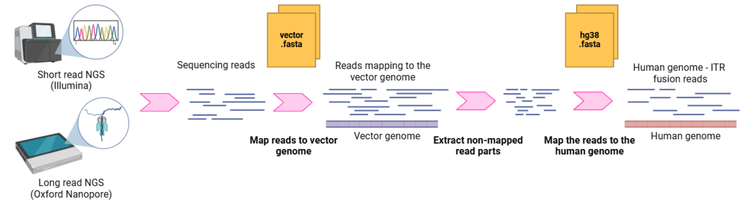
A distinct vector payload (host cell genome, plasmid backbone,...)



AND THE CHIMERIC

An undesired sequence fused to one or two ITRs

## The Ascend analysis pipeline to detect rare fusion events

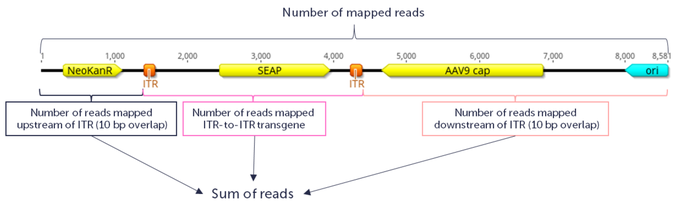



Depicted is the core pipeline for HCD-ITR fusions, however comparable strategies can be applied to detect any fusion event

- The pipeline employs subsequent mapping of reads to the vector and human genome and was implemented in an automated bash script
- It is agnostic to the read length and works with short and long reads

## ITR readthrough/bidirectional packaging cause vector plasmid backbone-ITR chimeras detected by long read NGS

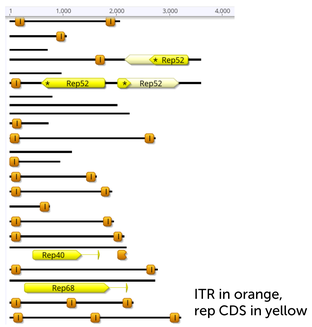
Vector/plasmid backbone chimeras = Sum of reads – Number of mapped reads



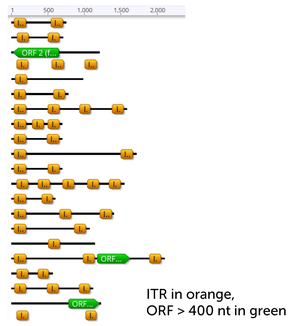

| Number of mapped reads | Sum of reads | Chimeric reads |
|---|---|---|
| 4,508,087 | 4,728,674 | 220,587 (≙ 4.89%) |

Reads were classified as chimeric if they extended more than 10 bp into the adjacent plasmid backbone

## Low rep-ITR and HCD-ITR fusion levels in the rAAV vector batch




Rep-ITR fusions are extremely rare events (0.029% of total reads). Most reads are too short to span full-length genes, i.e. only 79 rep-ITR fusions (0.0017% of total reads) contained at least one full-length rep gene.




HCD-ITR fusions are also detectable but represent very rare events (0.092% of total reads). The majority of reads is too short to encompass full-length human genes. The figures display nanopore sequencing data.

## Comparison of long read (Oxford nanopore) and short read (Illumina®) sequencing

| Long read sequencing (Oxford nanopore) | Short read sequencing (Illumina) |
|---|---|
| one read covers entire rAAV payload | rAAV payload fragmented in several reads |
| functionality/sizing of ITR fusions | functionality/sizing of ITR fusions remains unclear |
| full ITR read coverage | reduced ITR coverage (sequencing by synthesis) |
| vector/vector plasmid chimera analysis possible | reads too short for vector/vector plasmid chimera detection |
| sequencing depth of 3-5 million total reads per MinION flow cell | very high sequencing depth (several hundred million reads per flow cell) |

Nanopore sequencing has many advantages,
but if read depth is a concern Illumina sequencing is a suitable complement

## Summary

Our pipeline identified ITR fusions of all relevant DNA impurities, albeit at very different frequencies. The most common impurities arise from ITR readthrough/bidirectional packaging of vector plasmid backbone sequences. ITR fusions requiring a recombination event were much less frequent but reliably discovered using our proprietary bioinformatic analysis pipeline. The data displayed that most ITR fusions contained only partial genes which is a crucial finding for risk assessments and regulatory filings. Notably, whereas reducing the risk of oncogenicity, this does not alleviate the immunological potential of aberrant peptides from expressed gene fragments.

Whereas the sequencing depth of Illumina short read sequencing is unparalleled, the technology is impaired by reduced ITR coverage, and in addition, cannot resolve the size of the fused DNA sequences. This additional information can only be obtained using long read sequencing such as nanopore sequencing. Taken together, we present a pipeline for the discovery of chimeric reads derived from the vector plasmid as well as very rare recombination events. The acquired data can lower the risk of a substantial proportion of DNA impurities as ITR-less and characterize the remaining ITR bearing impurities in-depth.

Figures for this poster were partially created using bioRender (biorender.com).

**Aim higher**

Ascend Advanced Therapies | Fraunhoferstraße 9b | 82152 Planegg | Germany

www.ascend-adv.com
Florian.Sonntag@ascend-adv.com

Poster download
Password: AIMHIGHER