Recommendations for the Validation of rAAV Identity by Next Generation Sequencing

Jarrod Dean, Associate Director, Genomic Medicine Unit
BioAnalytics, NGS CMC
Sanofi

On behalf of BioPhorum ATMP Next Generation Sequencing workstream

USP Biologics Stakeholder Forum
Innovative Analytical Approaches CGT
Rockville, MD; 22FEB2024
Permission to use
The contents of this report may be used unaltered as long as the copyright is acknowledged appropriately with correct source citation, as follows ‘Entity, Author(s), Editor, Title, Location: Year’

Disclaimer
This document represents a consensus view, and as such it does not represent fully the internal policies of the contributing companies.

Neither BioPhorum nor any of the contributing companies accept any liability to any person arising from their use of this document.

The views and opinions contained herein are that of the individual authors and should not be attributed to the authors’ employers.

CONNECT COLLABORATE ACCELERATE
is a trademark of BioPhorum Operations Group.
Agenda

• Introduction to BioPhorum & ATMP Phorum
• Product Identity testing
• Recombinant Adeno Associated Virus (rAAV) Identity Testing
• Historical Methods for rAAV Genomic Identity Testing
• Next Generation Sequencing for rAAV Identity Testing

• BioPhorum White Paper Outline for NGS ID Testing of rAAV
  ➢ Initial considerations to Establish Internal NGS Workflows
  ➢ Sequencing Library Preparation
  ➢ Method Development Considerations: Sample Preparation
  ➢ Sequencing
  ➢ Bioinformatics (Analysis)
  ➢ Validation Plan Proposal
  ➢ Validation Plan Proposal (Specificity)
  ➢ Validation Plan Proposal (Instrument & Analysis)
  ➢ Long-read Sequencing Technologies
  ➢ Comparison of Sequencing Technologies for rAAVs and Use

• Conclusion
BioPhorum: a co-ordinated program of industry change

BioPhorum creates an environment where the global biopharmaceutical and device industry can collaborate and accelerate their rate of progress, for the benefit of all. We do this by:

- Bringing leaders together to create future visions that focus the industry’s energy on the key emerging opportunities
- Mobilizing communities of the top experts around these opportunities, up and down the biopharma value chain
- Creating partnerships that enable change and provide the quickest route to implementation and results
- Replacing isolation with collaboration so that the industry shares, learns and builds the best solutions together

...making the journey better, faster and cheaper than it would be for individual companies to do it on their own.

There are currently 11 Phorums providing a wealth of opportunities for companies to align their interests with similarly committed organizations.

150+ member companies
120+ global programs for change
7500+ leaders and subject matter experts
150 published papers, presentations and resources in the last 12 months
ADVANCED THERAPY MEDICINAL PRODUCTS (ATMP)

Formed in 2018, the ATMP Phorum supports the quest for better and faster development of Cell, Gene (in vivo & ex vivo) and RNA therapies.

In this collaboration, we connect drug development and contract manufacturing organizations with the aim of ensuring harmonization and alignment around many issues. These include potency assays, phase-appropriate guidance for critical quality attributes, operator safety, regulatory guidelines, and ATMP specific validation issues, as well as working towards resolving the current challenges to commercializing ATMP products.

The business of cell, gene and RNA therapies is as diverse as the patients which it serves and for that there is no one size fits all solution.

ATMP members have identified eleven high level areas of interest, along with additional specific challenges to resolve. Addressing these challenges will help to avoid multiple industry solutions.

There are now two types of workstreams:

- **High Level Workstreams** allowing opportunities for quick topic discussion across the industry
- **In-Depth Workstreams** which will focus on a specific focused topic working towards a defined deliverable.

- Currently 14 in-depth workstreams

HIGH LEVEL WORKSTREAMS

- Cell therapy (+/- gene modified)
- *in vivo* gene therapy
- RNA

- Development
- Analytics & assay validation
- Commercial Readiness & CMC
- Regulatory
- EHS & Biosafety
- Raw Materials
Product Identity Testing

• Product lot testing ensures the safety and efficacy of a therapeutic active pharmaceutical ingredient

• Identity testing is specific to each product, confirming that the product is what is stated on the final container/package label

• Must be able to distinguish from other products being manufactured / processed in the same laboratory

• Release method (i.e., qualified / validated)

• Identity is established in multiple Code of Federal Regulations (CFR) Title 21 parts/subchapters

Recombinant Adeno-associated Virus Identity Testing

Identity is typically measured from the 5'-ITR through the 3'-ITR with emphasis on the therapeutic transgene of interest.

ii. Identity

We recommend that identity assays uniquely identify a product and distinguish it from other products in the same facility. This test is performed on the final labeled product to verify its contents (21 CFR 610.14). Sometimes, a single test is not sufficient to distinguish clearly among products, and therefore, it is good practice to use different types of test methods (e.g., vector genome restriction digest and protein capsid analysis).

Historical Methods for rAAV Genomic Identity Testing

<table>
<thead>
<tr>
<th>Method</th>
<th>Strategy</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>qPCR/dPCR</td>
<td>Product-specific amplification</td>
<td>- Well established methodologies</td>
<td>- Might not span entire ITR-ITR sequence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Minimum instrumentation with IQ/OQ</td>
<td>- Multiple assays required for genome integrity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Use of product specific oligos</td>
<td>- PCR assay spanning complete ITR is difficult</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Multi-attribute with vector genome</td>
<td>- Difficulty amplifying secondary structures</td>
</tr>
<tr>
<td>Restriction Enzyme Digest with PCR</td>
<td>Product-specific amplification with sequence specificity to restriction enzyme cut sites</td>
<td>- Target majority of transgene</td>
<td>- Assay design more complex</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Added sequence specificity with RE sites</td>
<td>- Requires more instrumentation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Make unique fingerprint for each product</td>
<td>- Limited to number of samples per assay</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Compare pattern directly to reference</td>
<td></td>
</tr>
<tr>
<td>Sanger Sequencing</td>
<td>Direct sequencing of full transgene</td>
<td>- Full ITR - ITR sequence coverage</td>
<td>- Low sequence depth (read accuracy)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Not enough reads for variant analyses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Difficulty sequencing through ITRs</td>
</tr>
</tbody>
</table>

- Non-inclusive list

- Shortcomings with above historical methods
  - Inability to provide an ITR – ITR full length sequence coverage readout (q-/dPCR, RE)
  - Depth of coverage insufficient for robust variant analyses (Sanger)

How can Next Generation Sequencing Improve rAAV Identity Confirmation?
Next Generation Sequencing for rAAV Identity Testing

- Massively parallel sequencing technology from an unbiased, holistic perspective
- Drastically increase the per base coverage of the sample, providing greater coverage depth
  - Increase confidence of variant calling
- Growing interest in industry of using NGS as an analytical tool
  - Identity and other parameters for product characterization
- Approaches are being provided by CROs/CDMOs
  - Offering validated solutions for identity testing and confirmation
BioPhorum Whitepaper for NGS Identity Testing of rAAV

Objectives

- Build on recent industry interest and promote NGS as an analytical tool for rAAV testing
- Define NGS rAAV identity testing workflow using a modular approach via short-read sequencing
  - Sample preparation
  - Library preparation
  - Sequencing
  - BioInformatics / Analysis
- Considerations for method development and control strategies for each workflow module
  - Identify target system suitability and sample acceptance criteria to ensure quality results
- Considerations for method validation
- High-level overview of other technologies and how they can be used for rAAV characterization
- Whitepaper is currently in draft and includes input from 29 SMEs across 21 member companies
Initial Considerations for Internal NGS Workflows

- Segregation of Pre and Post PCR Activities
  - Avoid contamination during workflow
    - No template clean area (reagent preparation)
    - Pre-PCR clean area (sample preparation)
    - Post-PCR area (PCR and sequencing workflows)

- Lab Space and Utilities
  - Instrument location (i.e., vibration considerations, direct sunlight)
  - Lab temperature and humidity considerations
  - Uninterrupted power supply (UPS)
  - High speed network for large data management
  - Other considerations depending on sequencer (i.e., Nitrogen line)

- Instrument and Analysis Validation and GMP Compliance for Data Integrity and Storage
  - Adherence to cGMP guidance (i.e., 21 CFR Part 11 and Annex 11) for all workflow instruments
  - Validation plan for analysis application and environment
    - Requires input from matrixed team (i.e., Digital, IT and local business quality representatives)
  - Data storage considerations
  - Production environment support with established change control process

Engineering controls (i.e., dedicated BSC) in lab if space is limited
Method Development Considerations: Sample Preparation

Example Short-read SBS workflow

Sample Input (rAAV vector genome) → DNase Treatment → Proteinase K and Heat treatment → DNA Extraction → Second-strand Synthesis/Thermal Annealing → Second-strand synthesis Fragmentation → Quality Control

- Each step could be considered a critical assay attribute
  - Internal method development should be considered
  - Define robust range to ensure optimal method performance

- Establish control to monitor assay performance
  - Controls may not transfer over to final version of method
    - Spiked DNA control to show DNase treatment efficiency

- Suggested quality metrics for performance monitoring
  - Final extracted DNA
    - Yield and purity
  - DNA sizing (post fragmentation)
    - Fall within established size range

- Potential opportunity for universal reference standard
  - Should be run during each analysis as a quality control
Many commercial library kits available for sequencing library generation
- Recommended to perform due diligence and compare kits for performance check
- Recommended quality metrics
  - DNA library size distribution, purity, and concentration
- “One size fits all” method approach may not be appropriate when considering unique transgene configurations

Sequencing

- Sequencing run quality is directly related to sequencing library quality
  - Robust wet-lab procedure is required for high quality results
- Required results will inform sequencing platform and associated kit
  - Depth of coverage and/or number of samples, per run

- Illumina MiSeq

- Sequencing quality considerations
  - Cluster density range
  - Phred score (% Q30)
  - Reads passing filter
  - Number of reads per sample

Percent PhiX control spike consideration, depending on complexity of library, to monitor error rate
Multiple Tooling Options

<table>
<thead>
<tr>
<th>Step</th>
<th>Tool</th>
<th>Language</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adapter</td>
<td>Trimmomatic</td>
<td>Java</td>
<td>Widely used</td>
</tr>
<tr>
<td>Trimming</td>
<td>bbDuk</td>
<td>Java</td>
<td>K-mer based, part of bbTools</td>
</tr>
<tr>
<td></td>
<td>Cutadapt</td>
<td>Python</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>Trim Galore</td>
<td>Perl and</td>
<td>Wrapper for Cutadapt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Python</td>
<td></td>
</tr>
<tr>
<td>Assembly</td>
<td>SPAdes</td>
<td>C++</td>
<td>Small genomes, single-cell and metagenomics</td>
</tr>
<tr>
<td></td>
<td>Velvet</td>
<td>C</td>
<td>One of the first de-novo assemblers</td>
</tr>
<tr>
<td></td>
<td>ABYSS</td>
<td>C++</td>
<td>Large genomes</td>
</tr>
<tr>
<td>Mapping</td>
<td>BBMap</td>
<td>Java</td>
<td>Accurate, fast, high precision</td>
</tr>
<tr>
<td></td>
<td>BWA</td>
<td>C</td>
<td>Widely used for large genomes</td>
</tr>
<tr>
<td></td>
<td>Novoalign</td>
<td>C++</td>
<td>Commercial, accurate for a cost</td>
</tr>
<tr>
<td></td>
<td>Bowtie2</td>
<td>C++</td>
<td>Balances speed and accuracy</td>
</tr>
</tbody>
</table>

- Requires coding environment to support multiple scripting languages
  - Multiple programs should be evaluated during method development as they could impact results
  - Code development should be conducted using semantic version control to track and comment changes
- Final outputs could include alignment mapping (depth of reads), consensus sequence, and variant detection
- Quality criteria should be built around percent of reads aligning to transgene and depth coverage
ICH Q2(R2) Guidelines

NGS is a complicated assay

Will detect all sequences present in a sample
  • Common elements between rAAV vectors?

Reads are fragmented for short-read workflow
  • Not continuous ITR – ITR sequence detection

Best way to prove identity confirmation of a specific product?
Validation Plan Proposal (Specificity)

- Four (or more) different rAAV vectors and established control sequenced in same run
- The reference database will be a combination of the AAV transgene sequences
- Specificity
  - The reads from each rAAV will be aligned with the dataset
  - Therapeutic gene of interest (GOI) detected in corresponding reference only

![Sequence and align with references](Image)

- Must also include ITR – ITR confirmation of each AAV product in separate analyses
- Consider running formulation buffer / NC preparation to control reagent suitability
Validation Plan Proposal (Instrument and Analysis)

• Instrumentation
  • Validation handled by internal company standards (i.e., URS, SOP, etc.)
  • Review of vendor IQ/OQ documentation and generation of supplemental protocol (if necessary)

• Analysis Pipeline and Environment
  • Conforms to regulatory standards (i.e., 21CFR Part 11, Annex 11, etc.)
    • Guidance for Industry and FDA Staff: General Principles of Software Validation
  • Internal risk assessment based upon function and non-functional requirements
  • Internal code review of pipeline
    • Use of version-controlled processes (i.e., GIT)
  • Inclusion of audit trail for application / environment performance confirmation
  • Documented test cases in UAT environment
  • Wet-lab validation work helps inform suitability of dry-lab validation

• Robust validation package should include full description and functionality of the analysis application, environment architecture, UAT and wet-lab results
Long-Read Sequencing Technologies

Single-Molecule Real-Time (SMRT): PacBio

Nanopore Sequencing: Oxford Nanopore Technologies
Comparison of Sequencing Technologies for rAAV

<table>
<thead>
<tr>
<th>Short-read NGS</th>
<th>Long-Read NGS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Method</strong></td>
<td><strong>Method</strong></td>
</tr>
<tr>
<td>Advantages</td>
<td>Advantages</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Disadvantages</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Use Cases</td>
<td>Use Cases</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Combination of short- and long-read sequencing may be required for full analysis parameter understanding
- Robust method development must be conducted to understand impact of method and sequencing induced artifacts
Conclusions

• NGS can provide in-depth rAAV product understanding
  • Multiple sequencing technologies can provide enhanced insights to rAAV genomic diversity

• BioPhorum ATMP Next Generation Sequencing workstream is drafting a White Paper on the use of NGS for rAAV identity testing in a GMP setting

• There is expanding industry interest in establishing NGS as an analytical tool for rAAV
  • GMP services are being offered by external service providers
  • Updates to Regulatory Guidance are including the use of NGS [i.e., ICHQ5A(R2)]

• Current environment is supportive of utilizing NGS for rAAV release testing
  • Alignment with Regulatory Agencies through transparent communication on filings
    • Receive feedback from robust validation strategies and packages
    • Potential to establish universal controls to support method development and validation
Acknowledgements

**AbbVie**
Reshma Sundar

**Bayer**
Wenjie Yao
Jayashree Ray

**Charles River Laboratories, Inc.**
Georges Aboujaoudé

**Eurofins BioPharma Product Testing**
Weihong Wang

**FUJIFILM Diosynth Biotechnologies**
Andrew Hillhouse

**Merck & Co., Inc., Rahway, NJ, USA**
Meiping Chang

**Passage Bio**
Yuanyuan Gao

**PolyPlus Transfection**
Sylvain Julien

**Regeneron**
Aishwarya Shevade
David Van Houte

**REGENXBIO Inc.**
Bella Bidzhieva

**Roche**
Meenakshi Doma
Johannes Pschirer
Mehdi Sarmady

**Sangamo Therapeutics**
Rohit Rohra

**Sanofi**
Wei Zhang

**SK pharmaTECO Cell & Gene US**
Alexei Saveliev
Amanda Oran

**UCB**
Ishtiaq Khaliq

**Ultrasenx Pharmaceutical**
Wei Zhang

**uniQure biopharma B.V.**
Francesco Catania

**BioPhorum**
Niamh O’Kane
For further information about BioPhorum and the ATMP Phorum, please contact Steven Wall (ATMP Phorum Director)

Steven.wall@biophorum.com
Extra Slides
Short-read Sequencing Overview

Sequencing by Synthesis Approach (Illumina)

Sample and Library Preparation

Final Sequencing Library Configuration

DNA Sample -> Fragment DNA -> Damage repair -> End repair -> Ligate Adapters

For clustering: Libraries must have P5 and P7 binding regions on either end of a library
For sequencing: Libraries must have sequencing primer binding regions
For mixing samples: Libraries must have a unique index or barcodes sequence

Modified from: https://www.futurelearn.com/info/courses/a-practical-guide-for-sars-cov-2-whole-genome-sequencing/0/steps/338197
Short-read Sequencing Overview

Sequencing by Synthesis Approach (Illumina)

Data Analysis

- Demultiplexing
  - Sample specific indexes

- .Fastq file generation (per sample)
  - Contains all reads

- Used for subsequent analyses
  - Reference alignment
  - Visualization


Modified from: https://1010genome.com/illumina-sequencing-explained/