NGS Transcriptome Analysis in Cell Banking

Colette Côté
Chief Scientific & Portfolio Officer and US General Manager
FEB 2024
3.2.3: “NGS is encouraged as a replacement for in vivo assays because it can overcome the limitations of the breadth and sensitivity of virus detection of the in vivo assays. Furthermore this promotes the global objective to replace, reduce, and refine the use of animal testing.”

The guideline also refers to NGS as a replacement for:

- Mouse, Rat and Hamster Antibody Tests (3.2.4)
- PCR assays for virus specific detection. This can also help overcome the limitation of detection of virus variants (3.2.5)
- Targeted detection of known viruses based upon availability of their sequence (3.2.5.2).
- In vitro adventitious virus tests (3.2.5)
ICH Q5A(R2) clearly supports NGS as:

• Replacement for *in vivo* tests (supporting the 3Rs principle) for detection of inapparent viruses
• Replacement for HAP/RAP/MAP tests
• **Supplement** for 9CFR PCR panels
• **Supplement** for *in vitro* assays

- **Focus on cell-based systems** -
Transcriptome = analysis of expressed RNAs within cells
- Rapid & simple way to detect & identify any viral sequence in a sample
- Takes advantage of RNA phase of viral replication
  - All viruses go through an RNA phase during lifecycle
  - Minimizes background levels (host nucleic acid) making detection easier
- Detects all types of viruses
  - RNA/DNA
  - Circular/linear
  - Enveloped/non-enveloped
  - Single & double-stranded
- Analysis can be agnostic or targeted
  - Use strand info to assess active replication
  - Use in-process controls to assess matrix effects
Starting material is a cell line or purified RNA:
- In-process controls (negative and positive) processed in parallel
  - Neg = MRC5 cells; Pos = TPA-treated B95-8:Ramos (1:10^4)
- Use spiked-in collection of synthetic RNAs to track matrix effects
- Extract total RNA
- QC checks for quality/quantity (go/no-go gates)

Prepare a sequencing-compatible library:
- Convert material to dsDNA
- Add sequencing-compatible adapters
- Add tag to identify originating strand
- Remove rRNA to minimize background levels
- QC checks for quality/quantity (go/no-go gates)

Sequence using Illumina platform:
- NextSeq is the preferred instrument
- Multiple samples can be batched together

Analyze the data set:
- Proprietary algorithms (targeted and agnostic)
- Custom databases (viral/targeted)
- Output assessed for biological significance (secondary qualification & viral expert review)
Bioinformatics Basics

Agnostic Analysis

- Demultiplex data set
- Remove:
  - Short reads
  - Low quality reads
  - Homopolymers
  - Duplicates
  - Trim adapters
  - Remove host/rRNA

Assembly

- Assemble contigs
- Retain unassembled singletons

Unassigned Contig Evaluation

- Compare contigs & singletons against viral nucleotide DB
- Challenge against nt DB

Taxonomic Assignment

- Compare against viral protein DB
- Challenge against nt DB
- Challenge against protein DB

Automated Report

- Contigs & singletons that remain unclassified = unknown species
- If identified, flag as close or distant match, or background

Classify Hits

- Perform secondary qualification of the output for final determination
- Report

Unassigned Contig Evaluation

- Compare contigs & singletons against viral nucleotide DB
- Challenge against nt DB
- Challenge against protein DB

Automated Report

- Contigs & singletons that remain unclassified = unknown species
- If identified, flag as close or distant match, or background

Classify Hits

- Perform secondary qualification of the output for final determination
- Report

Targeted Analysis

Pre-processing

- Demultiplex data set
- Remove host/rRNA
- Remove:
  - Short reads
  - Low quality reads
  - Homopolymers
  - Duplicates
  - Trim adapters

Mapping

- Map reads against a custom database (clustered (compressed based on homology) or unclustered) containing defined reference sequences

Secondary Qualification

- Automated & manual assessments

Automated Report

- Perform secondary qualification of the output for final determination
- Report

Unassigned Contig Evaluation

- Compare contigs & singletons against viral nucleotide DB
- Challenge against nt DB
- Challenge against protein DB

Automated Report

- Contigs & singletons that remain unclassified = unknown species
- If identified, flag as close or distant match, or background

Classify Hits

- Perform secondary qualification of the output for final determination
- Report

Targeted Analysis

Agnostic Analysis
Secondary Qualification is Critical

- **Assay focuses on:**
  - Detection and strandedness of viral RNA sequences to assess presence of replicating viruses
  - Reduces impact of benign/free viral nucleic acids; minimizes false positives

- **Dedicated controls built into the assay to:**
  - Verify performance of the assay
  - Assess test sample matrix effects (e.g., inhibition)
  - Reduce false positives and false negatives

- **Expertise provides:**
  - Biological significance and context of any identified viral hits
  - Recommendations for mitigation/investigation if needed
Serial dilution of persistently infected B cell lymphocytes (B95-8 w/latent HHV4 (EBV)) in uninfected cells (Ramos)

Mimics low level infection

- Prepared 10X serial dilutions of B95-8 in Ramos (up to 1:10⁸)
- Cultured each in presence of TPA to express HHV4
- Counted cells & prepared fresh 1:10³ to 1:10⁸ dilutions in Ramos cells
- TPA treated B95-8 cells alone → positive control
- Untreated Ramos cells → negative control
- Sequenced & compared with HHV4 PCR results

NGS detected down to 1:10⁵ (same as PCR)
Comparability with *in vivo* and *in vitro* tests?

A quick note about spiking controls...

**Infected Cell Lines**
- Better reflects typical test matrix in the industry
- Reflects natural virus infection patterns & replication lifecycle
- Variable MOI but can prepare ratios of infected to non-infected cells to assess assay sensitivity/LOD

**Intact Purified Viral Particles**
- Broader spectrum of purified, characterized stocks available
- Quicker/simpler to use & most common spiking strategy
- Impact and kinetics of an active infection are lost
- Cannot be used to assess replication

**Nucleic Acids**
- Readily available stocks
- Easy to purify or synthesize & characterize
- Quick & simple to use
- Can be used to assess matrix effects
- Cannot be used to assess replication
Replacement for Animal Testing
ICH Q5A(R2): NGS is encouraged as a replacement for in vivo assays

Evaluation of a viral transcriptome Next Generation Sequencing assay as an alternative to animal assays for viral safety testing of cell substrates

Beurdeley et al., 2023: Head-to-head comparison providing supporting data for the industry
Replacement for in vivo HAP/MAP/RAP Testing
ICH Q5A(R2) refers to replacement with NGS

Mouse Antibody Production Test
- Ectromelia
- Mouse encephalomyelitis
- Lactate dehydrogenase elevating virus
- Hantaan
- Murine minute virus
- Mouse adenovirus
- Mouse hepatitis
- Pneumonia virus of mice
- Polyomavirus
- Sendai
- Epizootic diarrhea of infant mice
- Mouse cytomegalovirus
- Reovirus type 3
- Mouse pneumonitis virus
- Mouse thymic virus
- Mouse parvovirus

Hamster Antibody Production Test
- Sendai
- Pneumonia virus of mice
- Reovirus 3
- Lymphocytic choriomeningitis
- Simian virus 5

Rat Antibody Production Test
- Hantaan
- Kilham rat
- Mouse encephalomyelitis
- Pneumonia virus of mice
- Rat coronavirus
- Reovirus 3
- Sendai
- Sialoacryoadenitis
- Toolan

- Faster
- Animal-free
- More robust
- Fit for purpose
- Efficient

<table>
<thead>
<tr>
<th>Virus Family</th>
<th>Virus Species</th>
<th>Spikey CHO X® NK Cells (Copy Number)</th>
<th>Mouse Antibody Production Test</th>
<th>Hamster Antibody Production Test</th>
<th>Rat Antibody Production Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenoviridae</td>
<td>Mouse Adenovirus (MADV/MAV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Adenoviridae</td>
<td>Lymphocytic choriomeningitis virus (LCMV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Adenoviridae</td>
<td>Listeria monocytogenes meningitis virus (LMMV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Adenoviridae</td>
<td>Murine Polyomavirus (MPV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Sendai virus (SEDAV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Rat coronaviruses (RCAV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Harlan virus (HARV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Prosopis PB virus (PBPV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Mouse Cytomegalovirus (MCMV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Mouse Book virus (MBPV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Pneumonia virus of mice (PMV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Smallpox virus (SIV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Parvovirus B19</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Human parvovirus B19</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Mouse parvovirus (Mpv)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Retarda virus (RVDV)</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Sendai virus</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Mouse adenovirus</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Mouse hepatitis</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Pneumonia virus of mice</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Polyomavirus</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Sendai</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Epizootic diarrhea of infant mice</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Mouse cytomegalovirus</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Reovirus type 3</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Mouse pneumonitis virus</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Mouse thymic virus</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Mouse parvovirus</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
<td>Not detected</td>
</tr>
</tbody>
</table>

Control RNA ID | Negative Control | Positive Control | Acceptance Criteria | % Coverage | Ratio Sample/NK | Ratio Sample/PC | Conclusion |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>000150</td>
<td>60.44</td>
<td>99.94</td>
<td>% Coverage</td>
<td>60.98</td>
<td>112</td>
<td>112</td>
<td>PASS</td>
</tr>
<tr>
<td>000275</td>
<td>99.73</td>
<td>99.84</td>
<td>% Coverage</td>
<td>89.15</td>
<td>92</td>
<td>94</td>
<td>PASS</td>
</tr>
<tr>
<td>000353</td>
<td>82.85</td>
<td>78.58</td>
<td>% Coverage</td>
<td>81.25</td>
<td>96</td>
<td>910</td>
<td>PASS</td>
</tr>
<tr>
<td>000111</td>
<td>94.01</td>
<td>82.63</td>
<td>% Coverage</td>
<td>90.86</td>
<td>96</td>
<td>910</td>
<td>PASS</td>
</tr>
<tr>
<td>000200</td>
<td>89.73</td>
<td>96.76</td>
<td>% Coverage</td>
<td>84.32</td>
<td>94</td>
<td>97</td>
<td>PASS</td>
</tr>
<tr>
<td>000100</td>
<td>82.23</td>
<td>67.78</td>
<td>% Coverage</td>
<td>78.87</td>
<td>96</td>
<td>117</td>
<td>PASS</td>
</tr>
<tr>
<td>000055</td>
<td>78.78</td>
<td>82.70</td>
<td>% Coverage</td>
<td>96.90</td>
<td>122</td>
<td>116</td>
<td>PASS</td>
</tr>
</tbody>
</table>

The coverage of test item is 90% of the coverage of the NK or the PC for each Control ID.
“System Validation” broadly encompasses a range of method validations (e.g. extraction through data generation & reporting; including the software and its components (databases))

A validated system must be ‘fit for purpose’ -- in this case, must detect replicative adventitious viruses
Transcriptome Validation

Laboratory Method Validation:

- Mixture of infected: uninfected cells (EBV infected B95-8:Ramos or Raji:Ramos mixtures) spiked with known concentration (copy numbers) of a synthetic collection of representative target RNAs

- In-process controls included:
  - TPA treated B95-8 cells → EBV positive control
  - Raji cells → EBV positive control
  - Ramos cells → MuLV positive companion control
  - MRC-5 cells → Negative (uninfected) control

- Validation also included:
  - MRC-5 cells infected with Ad5, PI3, & Cox3
  - MDBK cells infected with BVDV

Bioinformatics/Software Validation:

- Used agnostic & targeted pipelines
- Screen for presence of EBV, MuLV, & spiked RNA controls
- Secondary qualification of the results by viral expert to verify hits
## Criteria for a Valid Test

<table>
<thead>
<tr>
<th>Method's step</th>
<th>Quality control</th>
<th>Quality acceptance criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA Extraction</td>
<td>Quality of extract (by Nanodrop):</td>
<td>Absorbance 200 / Absorbance 280 ≥ 1,8</td>
</tr>
<tr>
<td></td>
<td>Quality of RNAs (by Bioanalyzer):</td>
<td>- RIN (RNA Integrity Number) ≥ 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Major features of a successful ladder</td>
</tr>
<tr>
<td></td>
<td>Quantification of extract (by Qubit):</td>
<td>≥ 0,25 ng/μL</td>
</tr>
<tr>
<td>Library Preparation</td>
<td>Quantification of libraries (by Qubit):</td>
<td>≥ 3,00 ng/μL</td>
</tr>
<tr>
<td></td>
<td>Qualification of libraries (by Bioanalyzer):</td>
<td>Fragments range between 200bp and 1000bp with a local maximum at *≈*300–400bp</td>
</tr>
<tr>
<td>Sequencing</td>
<td>Illumina metrics:</td>
<td>- 65 million of reads at least per library</td>
</tr>
<tr>
<td></td>
<td>- Min read per library</td>
<td>- Q30 &gt; 80%</td>
</tr>
<tr>
<td></td>
<td>- Percent bases higher than Q30 (for flowcell high output 150 cycles)</td>
<td>- Chi2 test of homogeneity p-value &gt; 0,05</td>
</tr>
<tr>
<td></td>
<td>- Homogeneous repartition of read by libraries</td>
<td>- Output data at a minimum of 50gb</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>Negative control analysis (KNC-5 Cell Line [ATCC CCL-171])</td>
<td>- No replicative viruses are detected</td>
</tr>
<tr>
<td></td>
<td>Positive control analysis (TPA induced B95-8 cells [EBV infected] + Ramos cells [MuLV infected] ratio 1/10^9)</td>
<td>- Replicative HHV4 virus must be detected</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- MuLV virus genome has to be complete (&gt;95% coverage)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- No RNA control sequences identified as a positive viral hit</td>
</tr>
</tbody>
</table>
Enabling Regulatory Agency Support

ANSM review of PathoQuest’s Transcriptome Assay published 2023

<table>
<thead>
<tr>
<th>Modality</th>
<th>Process stage</th>
<th>Assays that can be substituted for PathoQuest NGS</th>
</tr>
</thead>
</table>
| Recombinant proteins      | • Unpurified bulk (with intact cells)  
• Master Cell Bank  
• Working Cell Bank  
• End of Production Cells | • *in vitro* adventitious agents  
• *in vivo* adventitious agents |

*"The wide detection spectrum of NGS assays allows the identification of viruses that cannot be detected by traditional assays, thus improving the diagnostic sensitivity of virological controls. This comparative study confirms the potential of the NGS transcriptomic assay as replacing traditional testing strategies, providing better assurance of safety and opening up new possibilities for safer innovative biologics and therapies."*

- Innovation Task Force (ITF)
- CBER Advanced Technologies Team (CATT)
- CDER Emerging Technologies Program (ETP)
- Drug or Biologics Master File
Cell Bank Characterization

Application of an NGS Transcriptome Assay

Production Systems for:
mAb’s, r-proteins, Vaccines, Gene Therapy, Cell Therapy

Master Cell Bank Characterization

**Microbiology**
- Sterility
- Mycoplasma

**Identity**
- Barcoding

**Adventitious Viruses**
- 3 Cell Line *in vitro*
- *In vivo*

**Retroviruses**
- Electron Microscopy
- Infectivity Assay
- F-PERT

**Species Specific Viruses**
- Mouse/Hamster/Rat Antibody Production
- Human Virus PCR panel

**Bovine and Porcine Viruses**
- Bovine Viruses
- Porcine Viruses

Faster, Safer, More Ethical Cell Bank Characterization

**Rapid Testing**
- Complete in as little as 2-3 weeks.
- Quicker release to manufacturing
- Faster to patients

**3Rs Compliant**
- Replacement, Reduction, Refinement
- In accordance with Directive 2010/63/EU

**Integrated Solution**
- Broad breadth and depth of coverage
- Replaces multiple assays
- Head-to-head comparisons completed
- Reduction in false positives
20 Years After Arriving on the Scene, NGS is Now:

- Robust, fit-for-purpose assay enabling targeted and agnostic approaches for contaminant detection
- Already integrated into guidance documents
- Supported by the agencies as a replacement assay for select classical tests (*in vivo*; HAP/MAP/RAP)
- Faster, safer, and animal-free testing
Thank you!

PathoQuest


contact@pathoquest.com