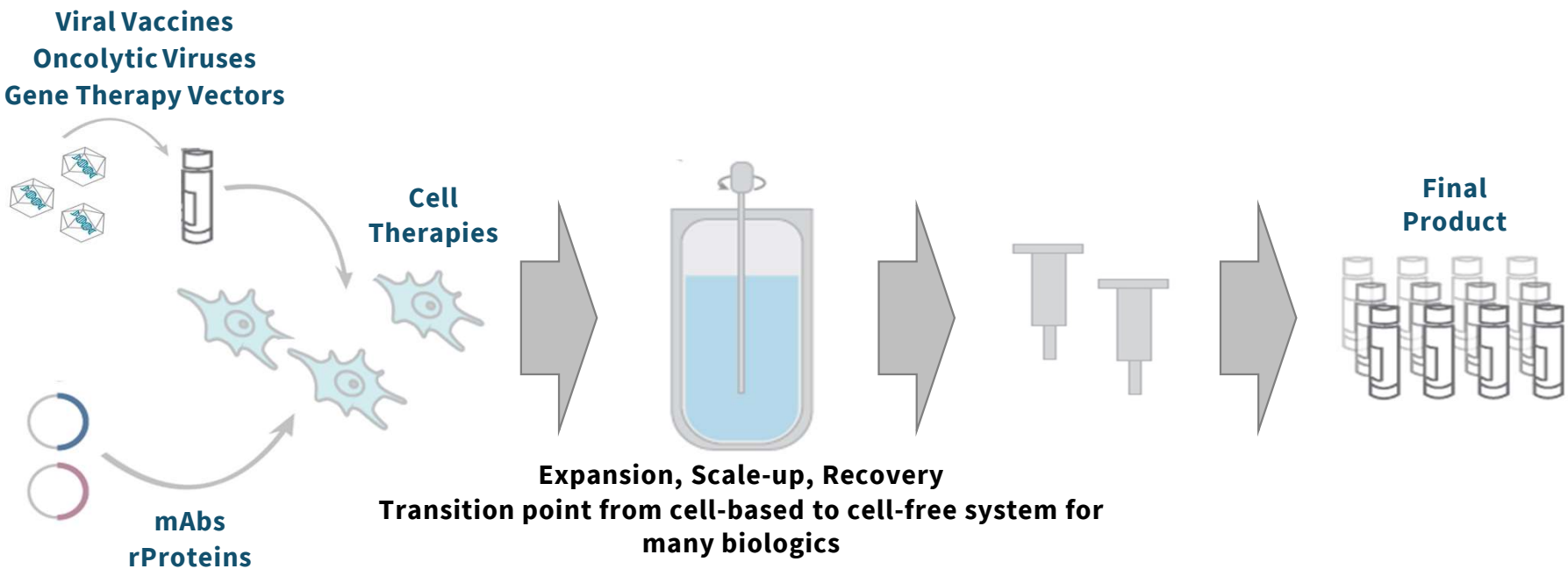
PathoQuest

NGS Process Overview

(Next Generation Sequencing for Biosafety Testing - A Primer)

September 2025
Colette Côté, Ph.D.
Chief Scientific and Portfolio Officer

Biologics Manufacturing At-A-Glance



Test Requirements:

Adventitious Agent Detection (Viral Safety) & Genetic Characterization

Traditional Biosafety Testing

Test Package Depends on Product & Regulatory Guidance

- Multiple, aging assays
- Includes animal-based testing
- Low resolution tests
- Extended timelines
- Increased cost

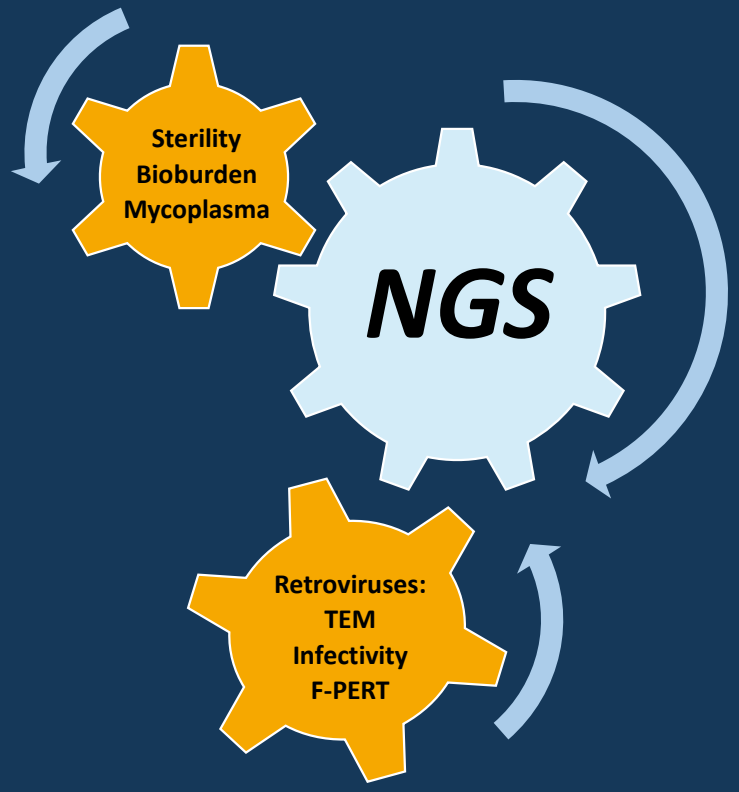
Classical Test Concerns:

- Resolution
- Specificity
- Lack of test system susceptibility
- Assay interference

Simplified Biosafety Testing

Goal:
One Test Package To Replace Several Assays

- Animal and quasi-animal based testing (*in vivo*, HAP/MAP/RAP)
- PCR tests
- *In vitro* tests
- Immunoassays (9CFR)



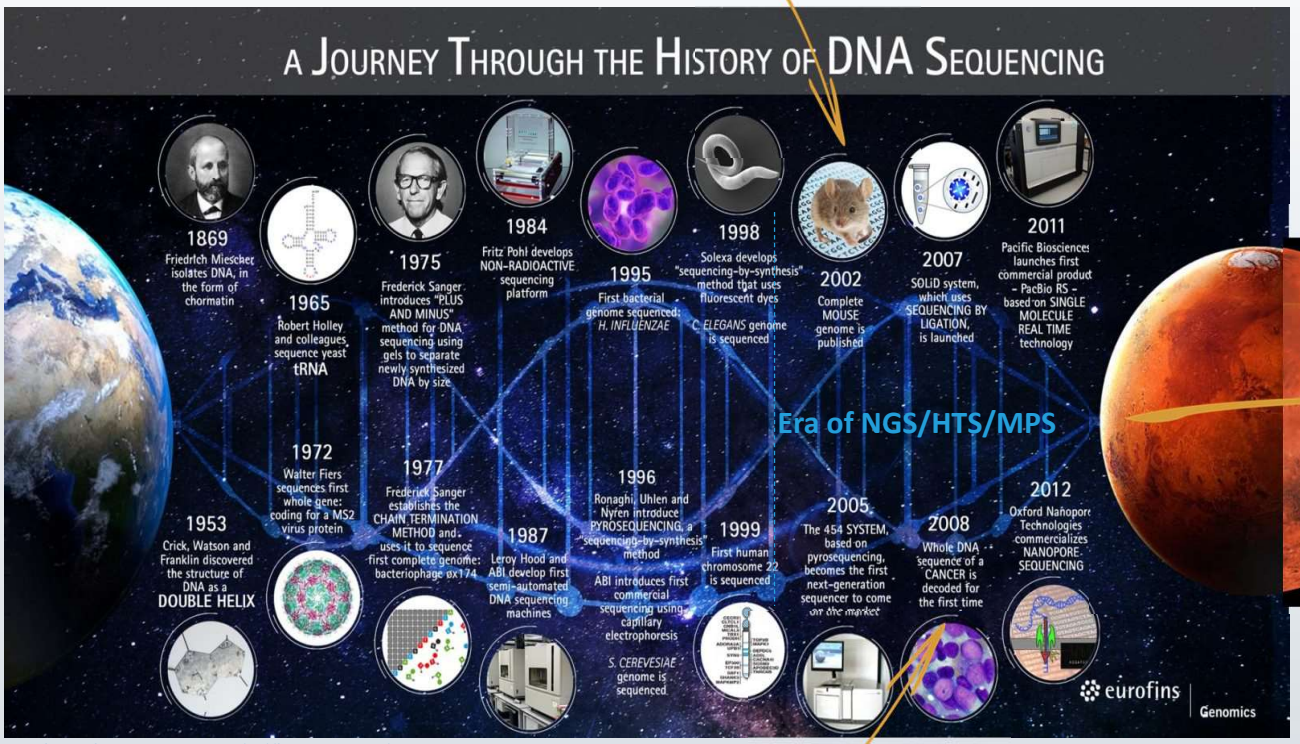
Advantages

- Improved resolution
- Improved specificity
- Improved sensitivity
- Simplified matrix interference assessments/method adoption
- No need for neutralizing antibodies or other pre-work

A Historical Foray Into Sequencing

2000

Lynx Therapeutics launched the first NGS technology called Massively parallel signature sequencing (MPSS). The company was later bought by Illumina.



A brief history of Next Generation Sequencing (NGS) (frontlinegenomics.com)
A journey through the history of DNA sequencing (the-dna-universe.com)

2008

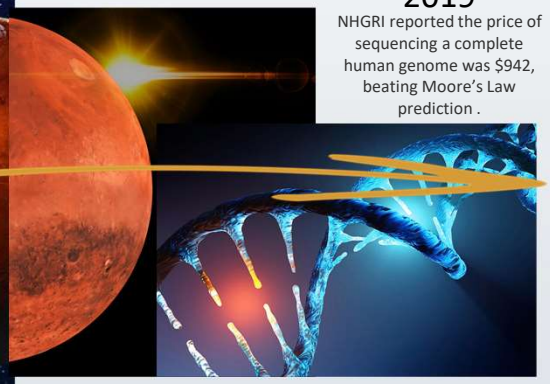
The first NGS human genome sequencing paper was published. James Watson's personal genome sequence was handed to him on a hard drive and was estimated to cost $1 million.

NGS enters the biosafety testing space

2014

Illumina launched the HiSeq X Ten Sequencer and the first $1,000 genome. Illumina held 70% of the market for DNA sequencers accounting for over 90% of all the DNA data produced globally.

First GMP NGS offered in the industry



2018

Veritas Genetics offered whole genome sequencing that cost just $199.

2019

NHGRI reported the price of sequencing a complete human genome was $942, beating Moore's Law prediction .

2022

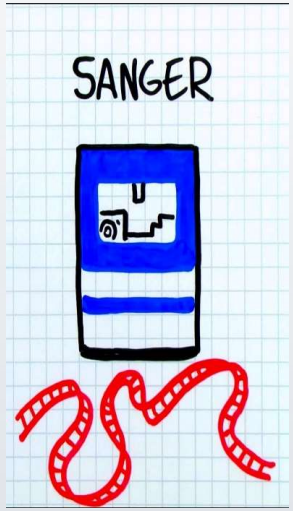
Ultima Genomics launches the $100 genome just 8 years after Illumina's HiSeq X Ten Sequencer breakthrough. Illumina counters it with the launch of NovaSeq X which promises to generate more than 20,000 whole genomes per year.

Today

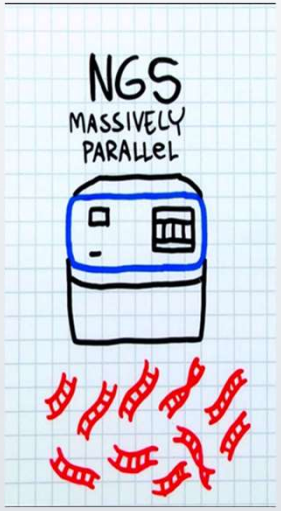
NGS explicitly included in international regulatory guidance & incorporated into regulatory filings

What's the Difference Between Sequencing Systems?

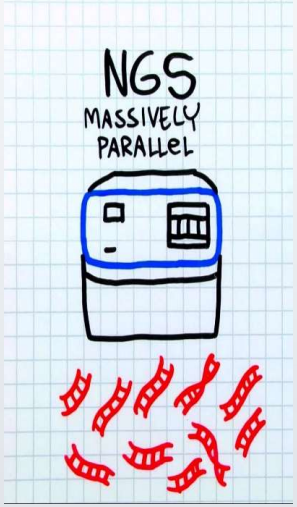
1st Gen
ABI 3730XL



2nd Gen
Roche/Illumina



3rd Gen
ONT/PacBio



Sequencing by Synthesis (SBS): Typically, < 1 kb or 'short read' sequencing

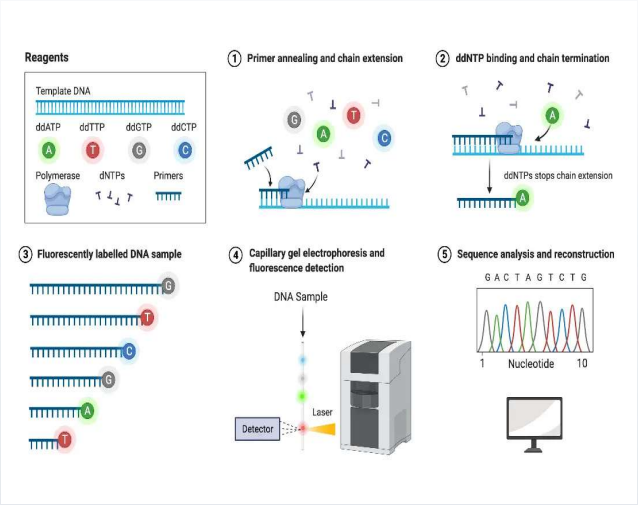
Nanopore Sequencing: 'Long read' or 'all read' sequencing up to ~200 kb

NGS is an advanced technology enabling analysis of *millions to billions* of individual DNA molecules rapidly & simultaneously
[Sanger is low throughput consensus sequencing]

NGS is sequence-agnostic—you don't need to know anything about the sequence to gain information about it
[Sanger is primer-dependent]

The Chemistry Explained

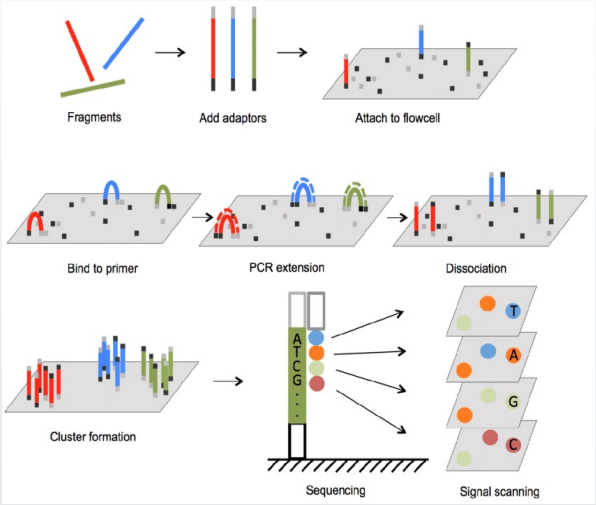
1st Gen (Sanger) Sequencing



SBS:

- Termination sequencing
- Determines consensus base for each position
- Individual forward and reverse reactions
- Primer-based
- Relatively low throughput

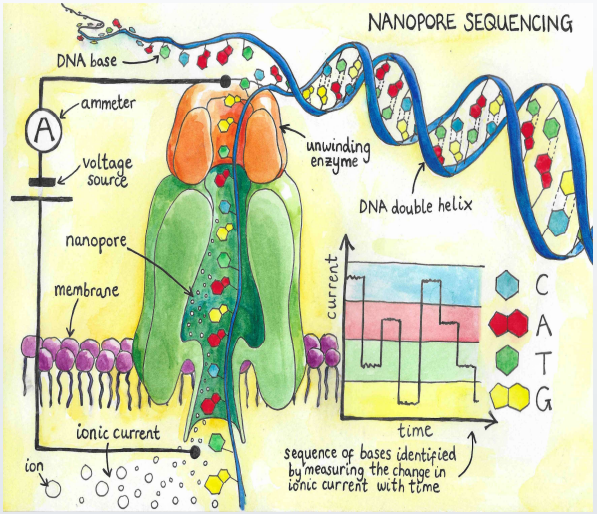
2nd Gen (Illumina) Sequencing



SBS:

- Base addition sequencing
- Color-coded bases
- Chemistry limiting (short reads)
- Individual read sequencing of both strands
- Very high throughput
- Multiplexing capabilities

3rd Gen (ONT) Sequencing

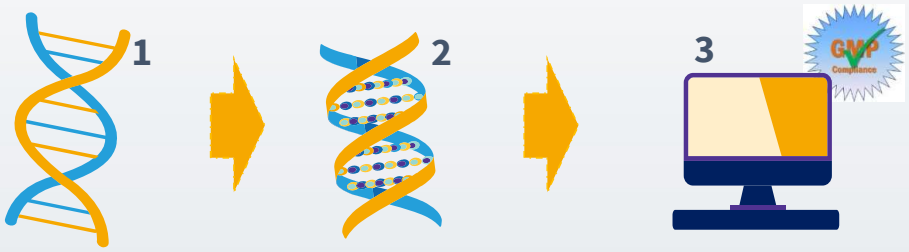


Nanopore:

- Direct sequencing of RNA or DNA strand moving through nanopore
- Detects current through pores & each base has a specific signature
- Not chemistry limited (ultra-long)
- Individual read sequencing of RNA
- Medium throughput

NGS Simplified

Three Basic Steps



Sample Processing (nucleic acid extraction/ library preparation)

Sequencing

Data Analysis/ Bioinformatics

Data Migration & Data Integrity Checks → **Data Pre-processing** → **Data Processing** → **Data Archiving/ Storage**

PATH: NGS Basics Sep 2025

Considerations:

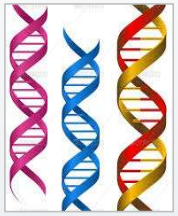
Critical Assay Elements:

- High quality sample processing = high quality data
- Smart bioinformatics (BFX)
- Fully validated, end-to-end GMP workflows
- Intermediate quality control steps (go/no-go gates)
- Appropriate assay controls

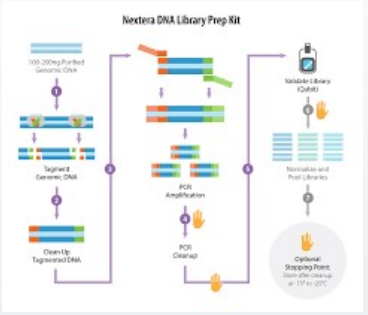
Bioinformatics

- Typically cloud-based computing, data storage & data transfer
- Scalable, dynamic environment
- Tailored BFX pipelines to meet current regulatory & client expectations
- Standardized databases/databanks for robust analysis
- Secondary qualification to minimize false positives

Diving Deeper into the Illumina Workflow



QC Checks



QC Checks

QC Checks

aws

Google Cloud

Microsoft Azure

Starting Material

- Wide range of starting materials (cells, virus banks, reagents, *etc.*)
- May require pre-processing (*e.g.*, (ultra)centrifugation, nuclease-treatment)

Nucleic Acid Extraction & QC Checks

- Extraction is process or sample type specific

Library Preparation & QC Checks

- Conversion of starting material to dsDNA for sequencing
- Adapter ligation/ addition of barcodes (indices) for clear sample identification & multiplexing

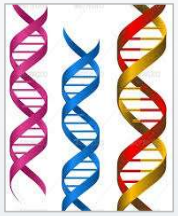
Sequencing & QC Checks

- Instrument preparation
- Sample pooling & dilution
- Sequencing
- Batching of samples is possible due to use of barcodes

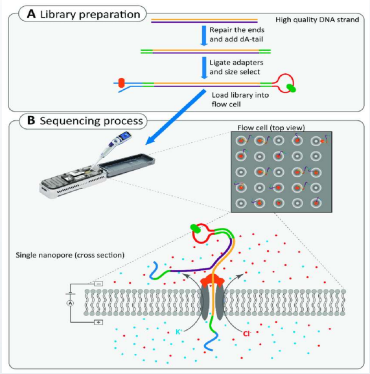
Data Analysis

- Typically cloud-based analytics & storage
- Custom & routine analysis
- Agnostic or targeted analysis
- Demultiplex samples for individual data set analysis

Diving Deeper into the ONT Workflow



QC Checks



QC Checks

QC Checks

aws

Google Cloud

Microsoft Azure

Starting Material

- Wide range of starting materials (cells, virus banks, reagents, *etc.*)
- May require pre-processing (*e.g.*, (ultra)centrifugation, nuclease-treatment)

Nucleic Acid Extraction & QC Checks

- Extraction is process or sample type specific
- HMW DNA is desired for many ONT applications

Library Preparation & QC Checks

- Simplified sample preparation
- Direct sequencing of RNA or DNA
- Base calling done in real time

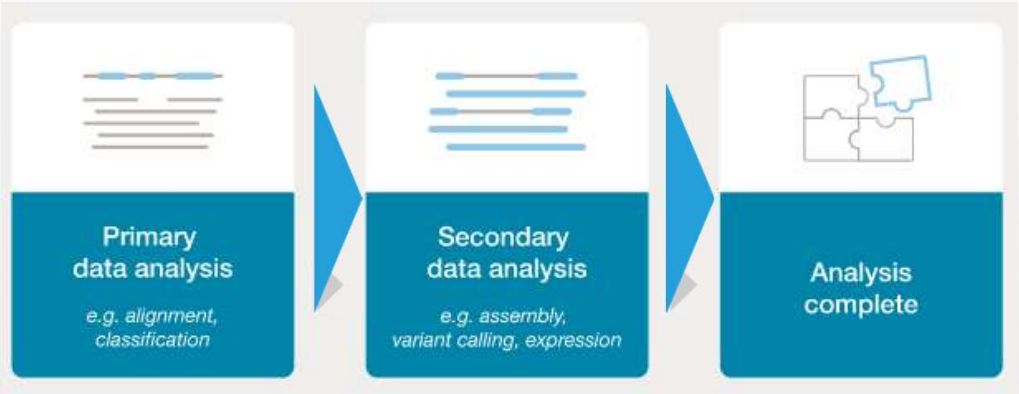
Sequencing & QC Checks

- Instrument preparation & flow cell check
- Sample pooling
- Sequencing
- Batching of samples is possible due to barcoding

Data Analysis

- Typically cloud-based analytics & storage
- Custom & routine analysis
- Agnostic or targeted analysis
- Demultiplex samples for individual data set analysis

Diving Deeper into Data Analysis



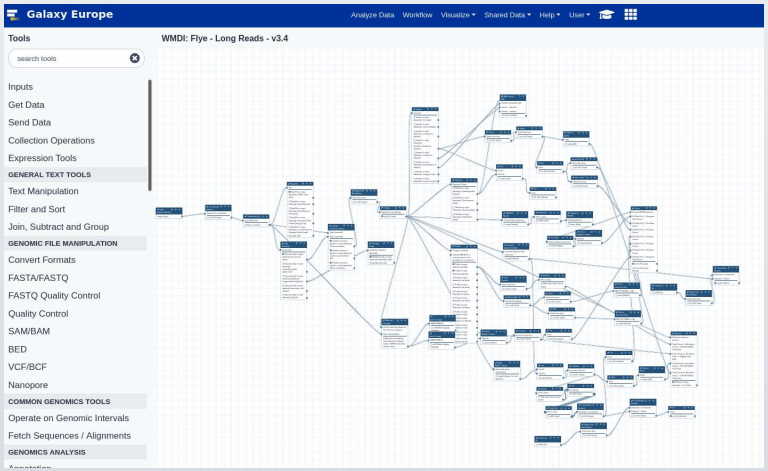
Starting Data Set

- Gigabytes to terabytes of data generated
- FASTQ or similar files
- Data integrity checks required
- Demultiplexing required if pooled samples

Quality Trimming/ Pre-processing

- Removal of poor-quality reads
- Trimming of ends
- Adapter removal
- Optional removal of other specific sequences

Multistep Data Processing



Equipment & Infrastructure Considerations

Sample Processing

- **Standard molecular biology equipment:**
  Required: vortex mixers, pipettors, centrifuges, thermocycler/heat blocks, fluorometer/ spectrophotometer, electrophoresis equipment
  Workflow specific: ultrasonicator (Covaris), ultracentrifuge, analytical balance
  Nice to haves (automation): nucleic acid extraction equipment (QIAGEN), liquid handler (PE, Tecan, Hamilton)

- **Equipment specific for quality control assessments (examples):**
  - **Fluorometer/spectrophotometer**
    - Qubit, Nanodrop, SpectraMax
  - **Electrophoresis equipment**
    - Bioanalyzer, Tapestation, Fragment Analyzer

- **Considerations:**
  - Throughput/scale
  - Equipment obsolescence
  - RUO vs 21 CFR Part 11/Annex 11 compliant
  - Integration with LIMS & other systems

Sequencing

- **Sequencer:**
  - **Long/all read:** ONT (MinION, GridION, PromethION), PacBio (Onso, Vega, Revio)
  - **Short read:** Illumina (MiSeq, NextSeq, NovaSeq), ABI Ion Torrent (Genestudio, Genexus)
  - Most QC is performed on-instrument

- **Considerations:**
  - Operating system
  - Throughput/scale
  - Required read lengths for analysis
  - Range of desired applications & supporting chemistries
  - Equipment obsolescence
  - RUO vs 21 CFR Part 11/Annex 11 compliant
  - Vendor lifecycle management
    - Software upgrades
    - Troubleshooting support

Data Analysis

- **Computing environment & supporting infrastructure:**
  - Cloud-based vs on-premise
- **Storage/archiving environment**
- **Data transfer/transit pathways & security checks**
- **Data analysis pipelines**
- **Supporting elements:**
  - Databases/databanks

- **Considerations:**
  - Operating system
  - Cloud provider & access to supporting validation documents
  - Open-source vs proprietary tools
  - Scalability & parallel processing
  - Lifecycle management
  - Data & infrastructure security
  - 21 CFR Part 11/Annex 11 compliance requirements
    - Data integrity checks
    - Electronic signatures
    - Access rights
    - Version control

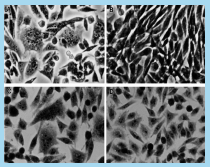
Other Considerations

Workflow Controls

- **Nature of the controls – are they suitable to the application?**
  - **In-parallel controls (positive and negative control samples)**
  - **Spiked-in controls (biological or synthetic, viral particles, nucleic acids)**

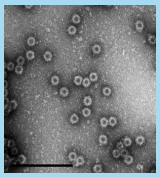
Infected Cell Lines

- 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 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



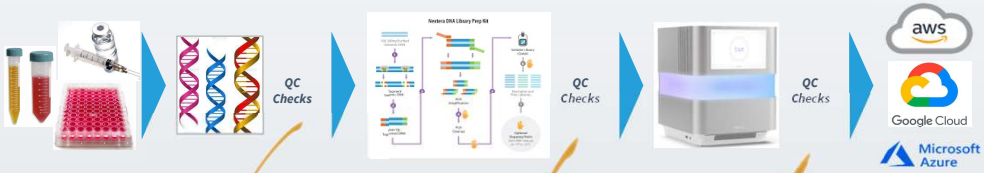
Nucleic Acids

- 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

- **Long term availability**
- **Level of characterization and available documentation**
- **Lot-to-lot consistency**

Quality Control Assessments

- **How many controls are needed?**
- **Where to place them?**
- **Nature of the assessment – quantitative/qualitative?**



Nucleic acid QC assessment

- Concentration
- Purity
- Integrity

Prepared library QC assessment

- Concentration
- Fragment size range

Sequencing run QC assessment

- Quality scores & other quality metrics (N50; cluster density)
- Total data output

Pipeline-specific QC assessments

Simplified Testing Strategy for Adventitious Agent Testing

Are Intact Cells Present?

Yes

Ex: MCB, WCB, EPOC, Crude Bulk Harvest

Transcriptomic Assay

- ‘Legacy’ assay in the field
- RNA-based test designed to detect replicating viruses
- Controls in place to evaluate assay performance & matrix effects

- Simplified, 3Rs compliant testing strategy
- Rapid TATs with broad detection capabilities
- Detects all types of replicating viruses
- Exquisite sensitivity, comparable to PCR
- Published head-to-head comparability data with classical testing

No

Ex: Clarified Bulk Harvest, DP, DS

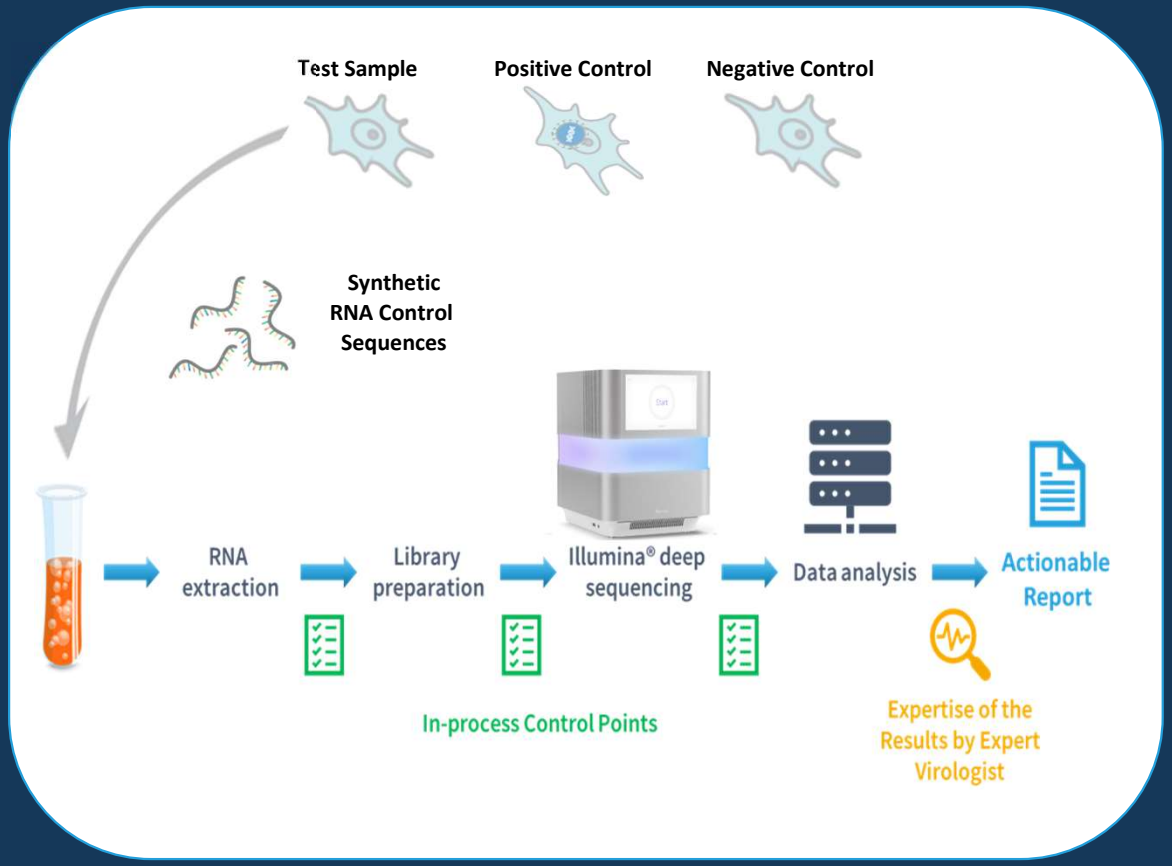
Viromic Assay

- Typically targets intact viral particles in cell-free systems
- No neutralization antibodies required

PATH: NGS Basics Sep 2025

Example: Transcriptomic Assay

Adventitious Agent Detection in Cell-based Materials



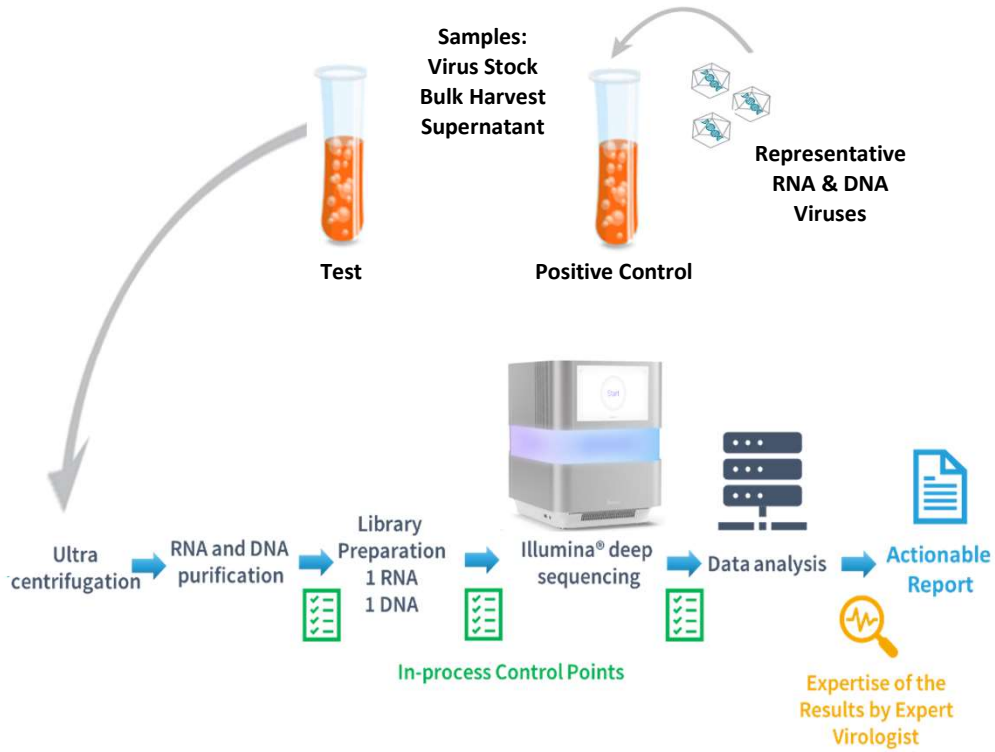
COMPARISON OF ADVENTITIOUS VIRUS DETECTION METHODS

	In vivo	In vitro	PCR	NGS
Speed	★☆☆☆☆	★★☆☆☆	★★★★☆	★★★★☆
Compatibility	★☆☆☆☆	★★☆☆☆	★★★★★	★★★★★
Detection of unknowns	★★★☆☆	★★★☆☆	☆☆☆☆☆	★★★★★
Identification of unknowns	☆☆☆☆☆	☆☆☆☆☆	☆☆☆☆☆	★★★★★
Volume required	★☆☆☆☆	★★☆☆☆	★★★☆☆	★★★★★
Detection of live replicating virus	★★★☆☆	★★★☆☆	☆☆☆☆☆	★★★★★

- **Fast Turnaround Time**
  - ***2-5 weeks; shorter than 28-day assay***
- **Low Input Sample Requirements**
  - ***Preserves more sample for other uses***
- **May use long or short read sequencing capabilities**
- **May include specific controls to satisfy method adoption requirements**
- **May use RNA strandedness to assess active replication**
- **Head to head comparability results published**
- **LOD comparable or better than PCR assays**

Example: Viromic Assay

Adventitious Agent Detection in Cell-free Materials



COMPARISON OF ADVENTITIOUS VIRUS DETECTION METHODS

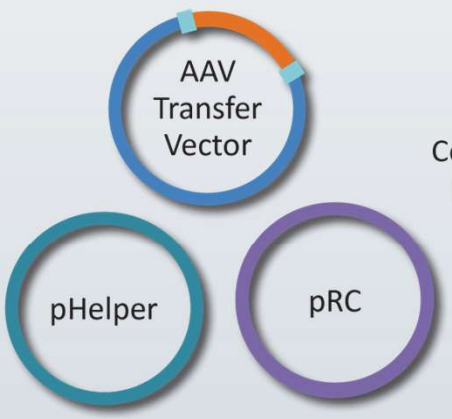
	In vivo	In vitro	PCR	NGS
Speed	★☆☆☆☆	★★☆☆☆	★★★★☆	★★★★☆
Compatibility	★☆☆☆☆	★★☆☆☆	★★★★★	★★★★★
Detection of unknowns	★★★☆☆	★★★☆☆	☆☆☆☆☆	★★★★★
Identification of unknowns	☆☆☆☆☆	☆☆☆☆☆	☆☆☆☆☆	★★★★★
Volume required	★☆☆☆☆	★★☆☆☆	★★★☆☆	★★★★★
Detection of live replicating virus	★★★☆☆	★★★☆☆	☆☆☆☆☆	★★★★★

- **Fast Turnaround Time**
  - ***3-6 weeks***
  - ***Shorter TAT than 28 day assay***
- **Low Input Sample requirements**
  - ***Preserves more sample for other uses***
- **May use long or short read sequencing capabilities**
- **Assesses all viral genome types present within viral particles**
- **LOD comparable or better than PCR assays**
- **Virus spiking studies used to assess matrix effects**

GMP GOOD MANUFACTURING PRACTICE CERTIFIED

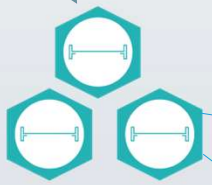
Other Applications in Vaccine Development

Identity Confirmation of Engineered Sequences



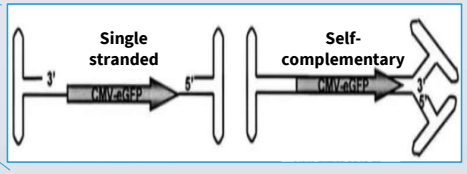
Replication Competency

Residuals



AAV viral vectors will be packaged and assembled.

Genome Integrity & Identity (Variants/Truncations/Flip flopping ITRs)



Simplified Testing Strategy for Genetic Characterization

Identity Confirmation

- Microbe-specific characterization
- End-to-end characterization in a single NGS run
- Exquisite variant detection (typically down to ~5% of the population)
- Purity determinations/sub-population characterization

- Rapid TATs with comprehensive characterization capabilities
- Exquisite sensitivity

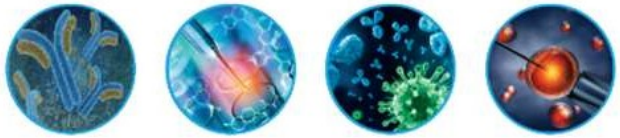
Integrity Assessment

- Evaluate intact versus truncated copies
- Evaluate encapsidated residuals
- Secondary identity confirmation

Thank you!

PathoQuest

Biologics Quality Testing. Faster, Safer.



contact@pathoquest.com