SANOFI PASTEUR

Sensitivity of a viral adventitious agent detection assay using HTS

Siemon Ng
Analytical R&D North America, Sanofi Pasteur

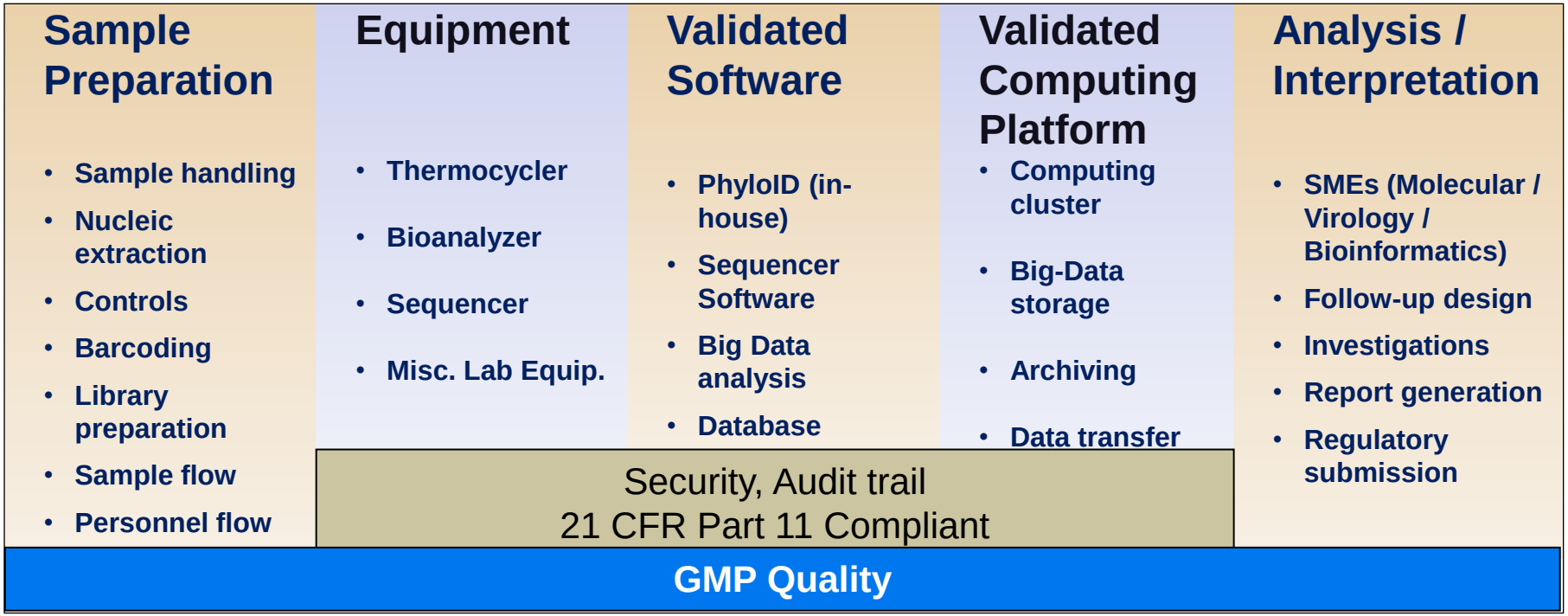
Outline

- **HTS Systems available at Sanofi Pasteur**
- **Validation approach for the sequencing**
- **Viral stocks for spiking studies**
- **Internal spiking studies**
- **Assay validation**
- **Regulatory perspectives**
- **Topics for discussion**

HTS Systems at Sanofi Pasteur Analytical Research and Development (North America)

- **Illumina HiSeq1500 (2013)**
  - Sequence 8 samples in ~11 days (validated)
- **Illumina NextSeq500 (2014)**
  - Single sample with rapid turnaround time (~3 days)
- **Illumina HiSeq3000 (Jan 2016)**
  - Sequence 8 samples in ~ 3 days
- **Ion S5 (2017)**
  - Including the Ion Chef (2015) that automates the library generation.

Detailed components of the validation approach

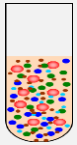


The Illumina HiSeq1500 was qualified in Oct 2016 for GMP use.

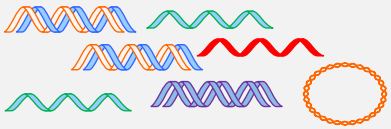
HTS Adventitious agent detection Assay Overview (1 / 2)

1. Sample Preparation

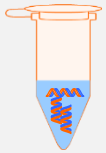
- Unbiased, efficient nucleic acid extraction and the creation of double-stranded DNA



Extraction



Second Strand Synthesis



Sample
(MCB, Harvest, DS etc.)

Nucleic Acid
(dsDNA, ssDNA, dsRNA, ssRNA, circular, etc.)

Double Stranded DNA

2. Sequencing Library Preparation

- Creating a sequencing library for Illumina HiSeq Sequencing

Enzymatic digestions to fragment the DNA

Add Illumina specific adaptors and barcodes

dsDNA

Fragmented DNA

Sequencing Library

HTS Adventitious agent detection Assay Overview (1 / 2)

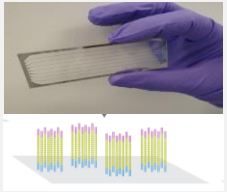
3. Sequencing

- Sequencing on the Illumina HiSeq1500

P5-Ind1-----DNA-----Ind2-P7

Sequencing Library

FlowCell Clustering



Cluster Generation

Sequencing

HiSeq 1500

Data Processing

```
ACTGGATCAG
CGATCTACGA
CGAGTTAGAT
TCCAGGATTC
AGGCTATTTC
GAGGATCTAG
```

Fastq Data File

4. Data Analysis

- Analysis on the PhyloID pipeline

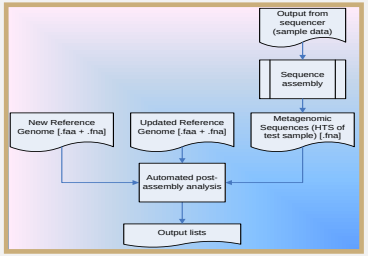
```
ACTGGATCAGCGAT
CTACGACGAGTTAG
ATTCCAGGATTCAG
GCTATTTCATTTCA
GGAAGCGACTTCTC
GAGGATCTAG
```

Fastq Data File

Data Transfer

Compute Cluster

Analysis



PhyloID

PASS

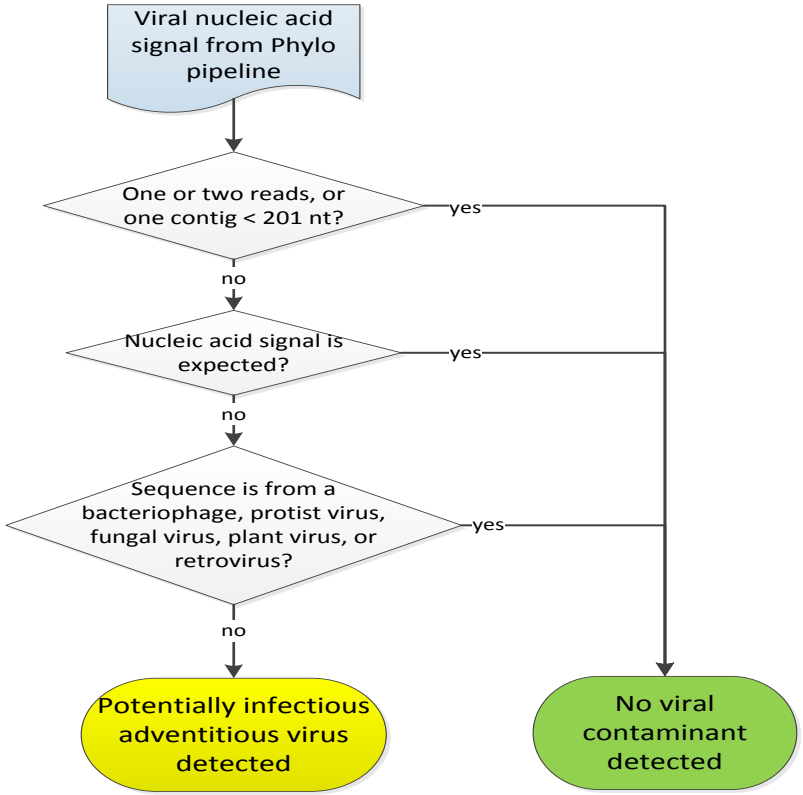
Follow-Up Investigation

Results

PhyloID™: An analysis pipeline developed in-house

- **2011: Started development of an automated analysis pipeline - PhyloID™**
  - Metagenomics approach where each read is identified
  - Big data challenge - the identification of low-abundance sequences from large scale sequencing data against the entire breadth of viral and organismal diversity
  - "Cataloguing the Taxonomic Origins of Sequences from a Heterogeneous Sample using Phylogenomics: Applications in Adventitious Agent Detection", R.L. Charlebois, S.H.S. Ng, L. Gisonni-Lex and L. Mallet, PDA J Pharm Sci and Tech, 68: 602-618, 2014
- **A sequence is identified based on its phylogenomic distance to a known set of reference sequences**
  - Both the match strength and the degree of fit can indicate confidence in the assignment.

What constitutes a positive signal?



SANOFI

Production of a viral stocks equivalent to the NIH viral stock

NIH study comparing the sensitivity between *in vivo* and *in vitro* assays

- **Link in-house generated HTS data to the *in vivo* published data from NIH**

Vaccine 32 (2014) 2916–2926

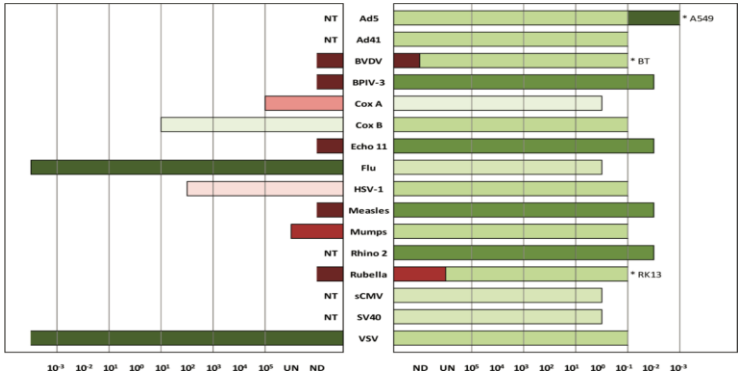
Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Systematic evaluation of *in vitro* and *in vivo* adventitious virus assays for the detection of viral contamination of cell banks and biological products☆

James Gombold[a], Stephen Karakasidis[a], Paula Niksa[b], John Podczasy[a], Kitti Neumann[a], James Richardson[c], Nandini Sane[c], Renita Johnson-Leva[c], Valerie Randolph[d], Jerald Sadoff[e], Phillip Minor[f], Alexander Schmidt[g], Paul Duncan[h], Rebecca L. Sheets[i,*]



Viral Family	Virus	Strain	Enveloped	Viral Genome	Genome Size (Kb)
Adenoviridae	Adenovirus 5	Adenoid 75	No	dsDNA	36
	Adenovirus 41	N/A	No	dsDNA	34
Flaviviridae	Bovine Viral Diarrhea Virus	NY-1	Yes	ssRNA (+ve)	12.4
Herpesviridae	Herpes Simplex Virus Type 1	MacIntyre	Yes	dsDNA	150
	Simian Cytomegalovirus	CS6	Yes	dsDNA	221
Orthomyxoviridae	Influenza A	A/PR/8/34 (H1N1)	Yes	8 ssRNA (-ve)	12.5
Paramyxoviridae	Mumps	Enders	Yes	ssRNA (-ve)	15.4
	Bovine Parainfluenza Type 3	N/A	Yes	ssRNA (-ve)	15.5
	Measles	Edmonston	Yes	ssRNA (-ve)	15.9
Picornaviridae	Coxsackie A16	N/A	No	ssRNA (+ve)	7.4
	Coxsackie B3	N/A	No	ssRNA (+ve)	7.4
	Echovirus 11	Gregory	No	ssRNA (+ve)	7.4
	Rhinovirus 2	HGP	No	ssRNA (+ve)	7.1
Polyomaviridae	Simian Virus 40	Pa-57	No	dsDNA	5.2
Rhabdoviridae	Vesicular Stomatitis Virus	Indiana	Yes	ssRNA (-ve)	11.2
Togaviridae	Rubella	M-33	Yes	ssRNA (+ve)	9.7

SANOFI

Internal Spiking Studies

Spiking into a viral harvest matrix

- The panel of 16 NIH viruses were spiked into a crude viral harvest at 10^3 to 10^5 genome copies per mL.
- All 16 viruses were detectable at a spike level of 10^5 and 10^4 genome copies per mL
- Coxsackie A16 was the only virus not detectable at 10^3 genome copies per mL
- Assay sensitivity was estimated to be approximately 10^3 genome copies for the 16 NIH viruses

Virus	Genome Copies (per mL)			
	10^5	10^4	10^3	10^2
Adenovirus 5	✓	✓	✓	✓
Adenovirus 41	✓	✓	✓	✓
Bovine Viral Diarrhea Virus	✓	✓	✓	✓
Herpes Simplex Virus Type 1	✓	✓	✓	✓
Simian Cytomegalovirus	✓	✓	✓	✓
Influenza A	✓	✓	✓	✓
Mumps	✓	✓	✓	✓
Bovine Parainfluenza Type 3	✓	✓	✓	✓
Measles	✓	✓	✓	✗
Coxsackie A16	✓	✓	✗	✗
Coxsackie B3	✓	✓	✓	✓
Echovirus 11	✓	✓	✓	✗
Rhinovirus 2	✓	✓	✓	✗
Simian Virus 40	✓	✓	✓	✗
Vesicular Stomatitis Virus	✓	✓	✓	✓
Rubella	✓	✓	✓	✗

SANOFI PASTEUR

Additional Replicates

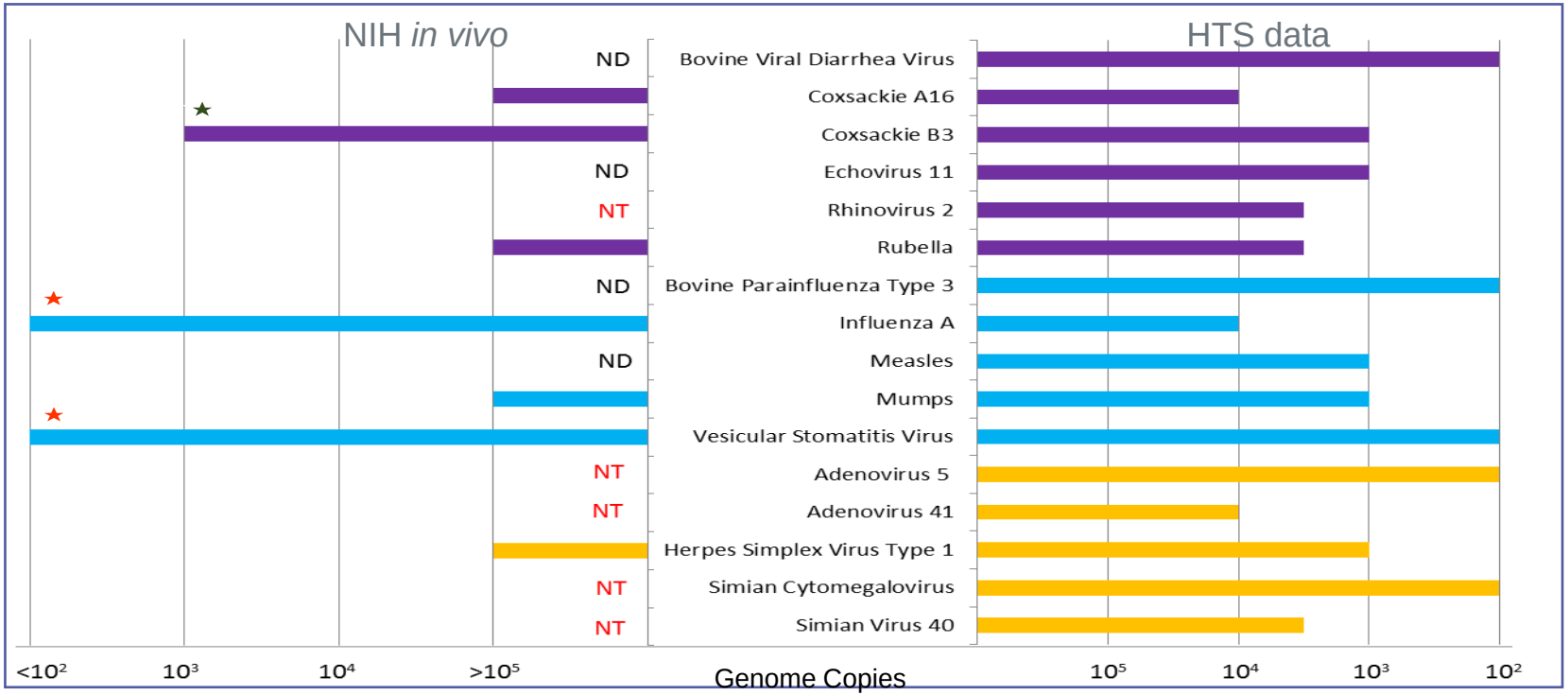
- Additional replicates using a similar spiking strategy was done.
- All 16 viruses were detected at 10^4 genome copies per mL, however some variability in the number of reads was observed at 10^3 genome copies per mL

- Assay specificity – Reads from some of the viruses were detected as a spread over 2 or 3 reference sequences (e.g. Coxsackie B3 sequences were also identified as Enterovirus B)
  - For this table, the identification of the virus was based on the strain with the highest number of reads.

Virus	Genome Copies (per mL)			
	10^4	10^4	10^3	10^3
Adenovirus 5	✓	✓	✓	✓
Adenovirus 41	✓	✓	✓	✗
Bovine Viral Diarrhea Virus	✓	✓	✓	✓
Herpes Simplex Virus Type 1	✓	✓	✓	✓
Simian Cytomegalovirus	✓	✓	✓	✓
Influenza A	✓	✓	✗	✗
Mumps	✓	✓	✓	✓
Bovine Parainfluenza Type 3	✓	✓	✓	✓
Measles	✓	✓	✓	✓
Coxsackie A16	✓	✓	✗	✓
Coxsackie B3	✓	✓	✓	✓
Echovirus 11	✓	✓	✓	✓
Rhinovirus 2	✓	✓	✗	✗
Simian Virus 40	✓	✓	✗	✓
Vesicular Stomatitis Virus	✓	✓	✓	✓
Rubella	✓	✓	✗	✓

Influenza A at 10^3 had ~50 reads in the previous study

Comparison of detection sensitivity between HTS and the *in vivo* assay



NT = not tested; ND = not detected;
Viral genome types: dsDNA; ssRNA (–ve); ssRNA (+ve)

★ *In vivo* assays show better sensitivity than HTS (and *in vitro* assays)

★ Equivalent between *in vivo* and HTS assay

Spiking into a Cell Bank matrix

- **Spiked the NIH model viruses with additional in-house viral stocks into a HEK293 cell matrix (1 million cells per mL)**

- **Spiked with known number of viral genomes**
  - 0.1, 0.01 and 0.001 genome copies per cell (10^5 , 10^4, 10^3 per mL)

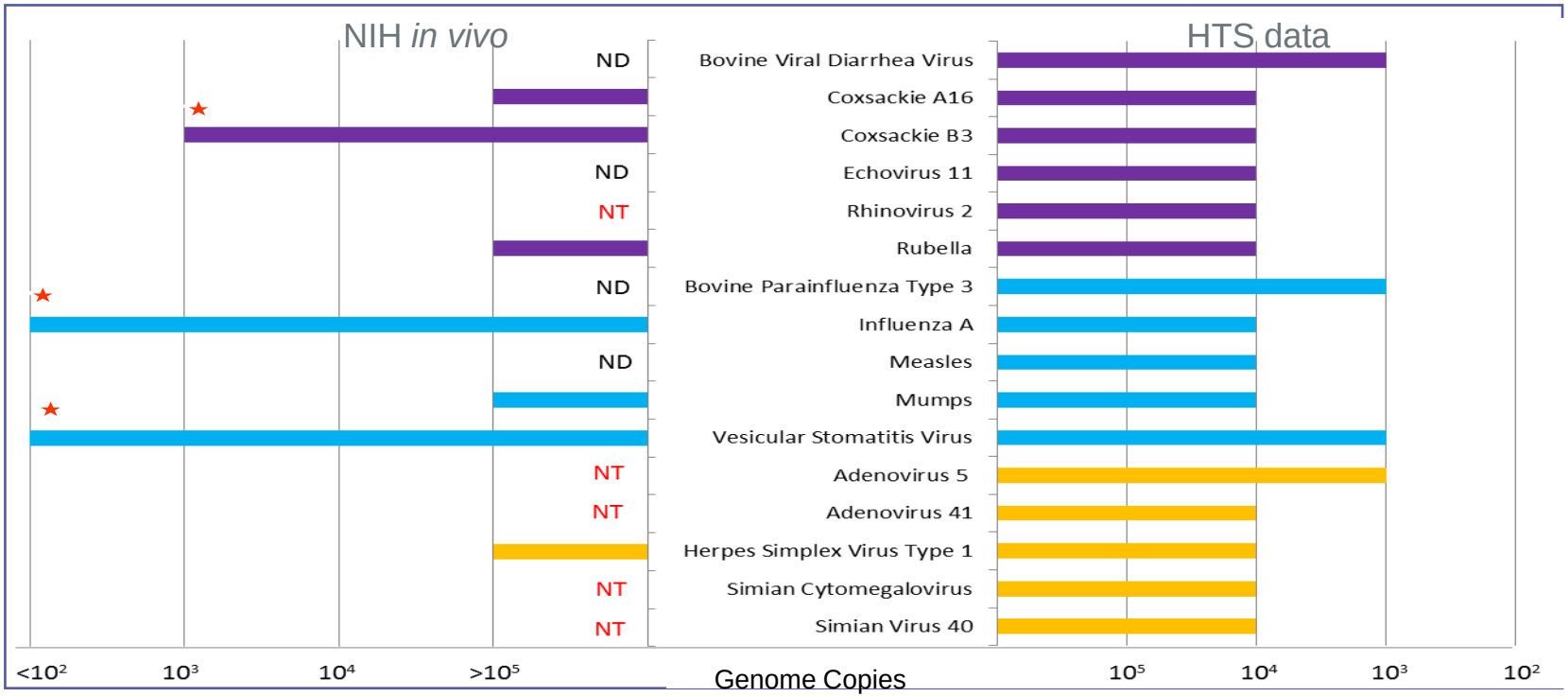
Cell bank matrix spiking study results

- Both replicates must be positive to be considered detected.
- 15 of the 16 viruses were detected at 0.01 genome copies per cell (10^4 copies per mL)
- Sensitivity for Influenza A is between 0.1 to 0.01 copies per cell based on a separate spiking experiment

Virus	Genome Copies per Cell			
	0.01	0.01	0.001	0.001
Adenovirus 5				
Adenovirus 41				
Bovine Viral Diarrhea Virus				
Herpes Simplex Virus Type 1				
Simian Cytomegalovirus				
Influenza A				
Mumps				
Bovine Parainfluenza Type 3				
Measles				
Coxsackie A16				
Coxsackie B3				
Echovirus 11				
Rhinovirus 2				
Simian Virus 40				
Vesicular Stomatitis Virus				
Rubella				

SANOFI PASTEUR

Comparison of detection sensitivity between HTS and the *in vivo* assay



NT = not tested; ND = not detected;
Viral genome types: dsDNA; ssRNA (–ve); ssRNA (+ve)

★ *In vivo* assays show better sensitivity than HTS (and *in vitro* assays)

Approach for assay validation

- Viral Adventitious Agent Detection by High-Throughput Sequencing is a complex assay involving multiple laboratory steps (with different kits, separate rooms and controls). We divided each of the major steps into their own procedure:
  1. Nucleic acid extraction and preparation
  2. Sequencing library preparation
  3. Operation of the sequencer
  4. Data analysis using PhyloID
  5. Putative positive adventitious agent follow-up assessment
- One assay that is applicable to cell banks, viral seed lots, bulk harvests and potentially any other biological samples.
- Dedicated rooms and biosafety cabinets with control of personnel flow and sample flow similar to that of a PCR assay.
- Assay will be qualified as a Limit of Detection assay
  - No quantitative analysis at this point.

Photo credits: ©

SANOFI

Topics for further regulatory discussion

Regulatory Perspective

- Evolution of regulations: Increasing acceptance of analytical tests for adventitious agent detection using new technology such as HTS
  - WHO TRS 978 Annex 3 (Recommendations for Cell Substrates):
  - Ph. Eur. Chapter 5.2.3, version 9.3: Cell Substrates for the production of vaccines for human use
  - Ph. Eur. Chapter 2.6.16, version 9.4: Tests for extraneous agents in viral vaccines for human use
  - Ph. Eur. Chapter 5.2.14, version 9.3: Substitution of *in vivo* method(s) by *in vitro* method(s) for the quality control of vaccines
- Waiving of *in vivo* tests
- How to evaluate qualitative elements (breadth of detection/specificity) and quantitative elements (limit of detection) to be comparable.

Points for further discussion

- Considerations for assay controls:
  - Negative Control(s)
  - Positive Control (a viral spike)
  - Extraction Control
- Some viral gaps might still exist based on the limited *in vivo* data available. A risk assessment will be performed to address any gaps in the validation data.
- Sensitivity / detection limit could potentially be matrix-dependent. Following a generic validation, a subset of the viral spikes may be used to do the matrix verification.
- Considerations for data submission:
  - Report
  - Raw data / metadata
  - Mechanism for data submission
  - Discussions at other conferences (e.g. HTS Standardization) regarding the use of Biocompute objects for data submission

Conclusions and Perspective

- Qualified a sequencing system under GMP environment
- Generated and characterized a panel of 16 viruses equivalent to those used in the NIH study comparing between *in vivo* and *in vitro* assays for use as viral spikes.
- Determined the sensitivity for using HTS for viral adventitious agent detection for both cell banks and viral harvests.
- Opportunity to streamline our adventitious agent testing package with the potential removal / supplementation / replacement / substitution of *in vivo* adventitious agent tests

Acknowledgments

- **Sanofi Pasteur**
  - Eric Abachin
  - Ali Azizi
  - Robert Charlebois
  - Virginie Courtois
  - Hugues Graf
  - Lucy Gisonni-Lex
  - Laurent Mallet
  - Rebecca Malott
  - Bryan McNeil
  - Sarmitha Sathiamoorthy
  - Martha Schreiber
  - Patrice Riou
  - Reza Yacoob
  - Anne-Gaëlle Darmont
  - Gee Marsh
  - Joan Bevilacqua
  - Clotilde Thiriart
  - Tony D'Amore
  - Marie-José Quentin-Millet
- **NIH: Dr Rebecca Sheets**
- **QBI and the Vaccine Research Program, Division of AIDS, NIAID:**
  - The viral strains were obtained through Quality Biological, Inc. and the Vaccine Research Program, Division of AIDS, NIAID, NIH: supplied under Contract No. HHSN272201100023C

THANK YOU

SANOFI