Detection of Adventitious Agents Using Next Generation Sequencing Technology

AFS ATCC FEDERAL SOLUTIONS

ATCC® Credible leads to Incredible™

Jeffrey I. Tokman[1], David Yarmosh[1], Joseph Petrone[1], Ana Fernandes[1], Sophia Huff[1], Kumari Meenu Karna[1], Anjan Panthee[1], John Bagnoli[1], Heather Couch[1]

[1]American Type Culture Collection, Manassas, VA

INTRODUCTION

The possible introduction of adventitious agents present a significant challenge in the production of biological products. Traditional detection methods, such as PCR-based techniques and live animal models, have notable limitations. PCR is restricted to known organisms, while live animal models are both time-consuming and costly. Moreover, the use of live animals is increasingly considered unacceptable due to ethical concerns.

Next Generation Sequencing (NGS) technology offers a comprehensive analytical method for detecting adventitious agents. Our research focuses on developing an NGS-based approach for this purpose. This method employs metagenomic classification using a minimally redundant database, multi-organism assembly, annotation, and quantification.

Furthermore, we incorporate differential expression analysis relative to a control sample to highlight any adventitious agent signals and reduce erroneous assignments. By utilizing various model viruses, we have established detection limits across different virus types, regardless of genome size and type (RNA or DNA).

MATERIALS AND METHODS

Virus	Strain	ATCC Number	Genome Type	Genome Structure	Genome Size	Envelope
Porcine circovirus (PCV1)	Type 1	FSCUST-91	ssDNA	Circular Genome	~1.7 kb	Non-enveloped
Respiratory syncytial virus (RSV)	A2	FSCUST-95	ssRNA (-)	Linear Genome	~15.2 kb	Enveloped
Epstein-Barr virus (EBV)	B95-8	FSCUST-94	dsDNA	Linear Genome	~180kb	Enveloped
Influenza Virus (Flu)	A/New Caledonia/71/2014 H3N2	FR-1400	ssRNA (-)	Linear Genome	~13.5kb	Enveloped

- **Cell Line:** HeLa cells, ATCC product number CCL-2, were as a background matrix for all spike studies currently tested.
- **Spike Experiment:** HeLa cells at 1x10^6 cells were spiked with a purified virus in triplicate at varying concentrations (Table 1), HeLa cells and media with no added virus was used as a control. Nucleic acid from the spiked samples was then extracted before sequencing and analysis.
- **Analysis:** The Kraken2 classification system was used along with a custom minimally-redundant database combined with the ATCC Genome Portal reference sequence. Reads were probed to determine the viruses that were potentially present in each sample. Reference sequences for each virus were used in gene-level quantification of each sample's reads through the Salmon tool alongside reads from a control sample. This allowed differential expression to separate the adventitious agent signal from the NGS noise.

Number of Samples	Spike Target	Host Cell	Library Type	RunID	Titers	Quantification Method
15	RSV (VR-1540)	HeLa	WGS	ILM1_2023AK	0.5, 1, 2.5, 5, 10 ng/uL	Qubit
9	RSV (VR-1540)	HeLa	WGS	ILM3_2023AZ	0.1, 0.25, 0.5 ng/uL	Qubit
12	PCV1	HeLa	WGS	ILM3_2023BK	0.1, 0.25, 0.5, 1, 2.5, 5 ng/uL	qPCR
4	PCV1	HeLa	WGS	ILM3_2023BL	0.1, 0.25, 0.5, 1 ng/uL	qPCR
9	RSV (VR-1540)	HeLa	WGS	ILN2_2023AR	0.1 ,3, 100 copies/cell	qPCR
9	RSV (VR-1540)	HeLa	WGS	ILN2_2023AU	0.1, 0.5, 5 ng/uL	
12	RSV (VR-1540)	HeLa	RNA-Seq	ILN2_2023AZ	0.1, 3, 100,1000 copies/cell	qPCR
13	RSV (VR-1540)	HeLa	RNA-Seq	ILN2_2024AF	(hostX1),0.01, 0.1, 3, 25 copies/cell	qPCR
13	PCV1	HeLa	RNA-Seq	ILN2_2024AM	(hostX1),0.01, 0.1, 3, 25 copies/cell	qPCR
13	PCV1	HeLa	RNA-Seq	ILN2_2024AP	(hostX1), 3, 25, 50, 100 copies/cell	qPCR
16	EBV	HeLa	RNA-Seq	ILN1_2024AR	(hostX1),0.001,0.1,1,3,1 copies/cell	
16	Flu	HeLa	RNA-Seq	ILN1_2024AR	(hostX1),0.001,0.1,1,3,1 copies/cell	

Table 1. Summary of spike experiments performed. Earlier runs WGS was used but later switched to mRNA-Seq to select for coding RNA not just background resulting in a greater signal/noise ratio.

RESULTS

Replicate	PCV Titer	Virus	ATCC Read Pairs	CLC Read Count	ATCC Genes Detected	Total Genes
JIT	3	PCV	1	0	2	2
KMK	3	PCV	0	0	0	2
SCH	3	PCV	1	0	2	2
JIT	25	PCV	15	0	2	2
KMK	25	PCV	14	16	2	2
SCH	25	PCV	17	0	2	2
JIT	50	PCV	32	10	2	2
KMK	50	PCV	33	8	2	2
SCH	50	PCV	39	12	2	2
JIT	100	PCV	76	12	2	2
KMK	100	PCV	69	10	2	2
SCH	100	PCV	84	22	2	2
Host	0	N/A	0	0	0	N/A

Table 2. LOD determination of PCV1 in HeLa cells at titers of 3-100 genome copies/cell. The ATCC Read pairs column indicates how many forward and reverse read pairs were classified using Kraken2 database for porcine circovirus, forward and reverse reads are counted independently. In the CLC read count column it can be seen at 25 gc/cell 2/3 samples were unable to make any identifications and no reads were identified in the 3gc/cell samples.

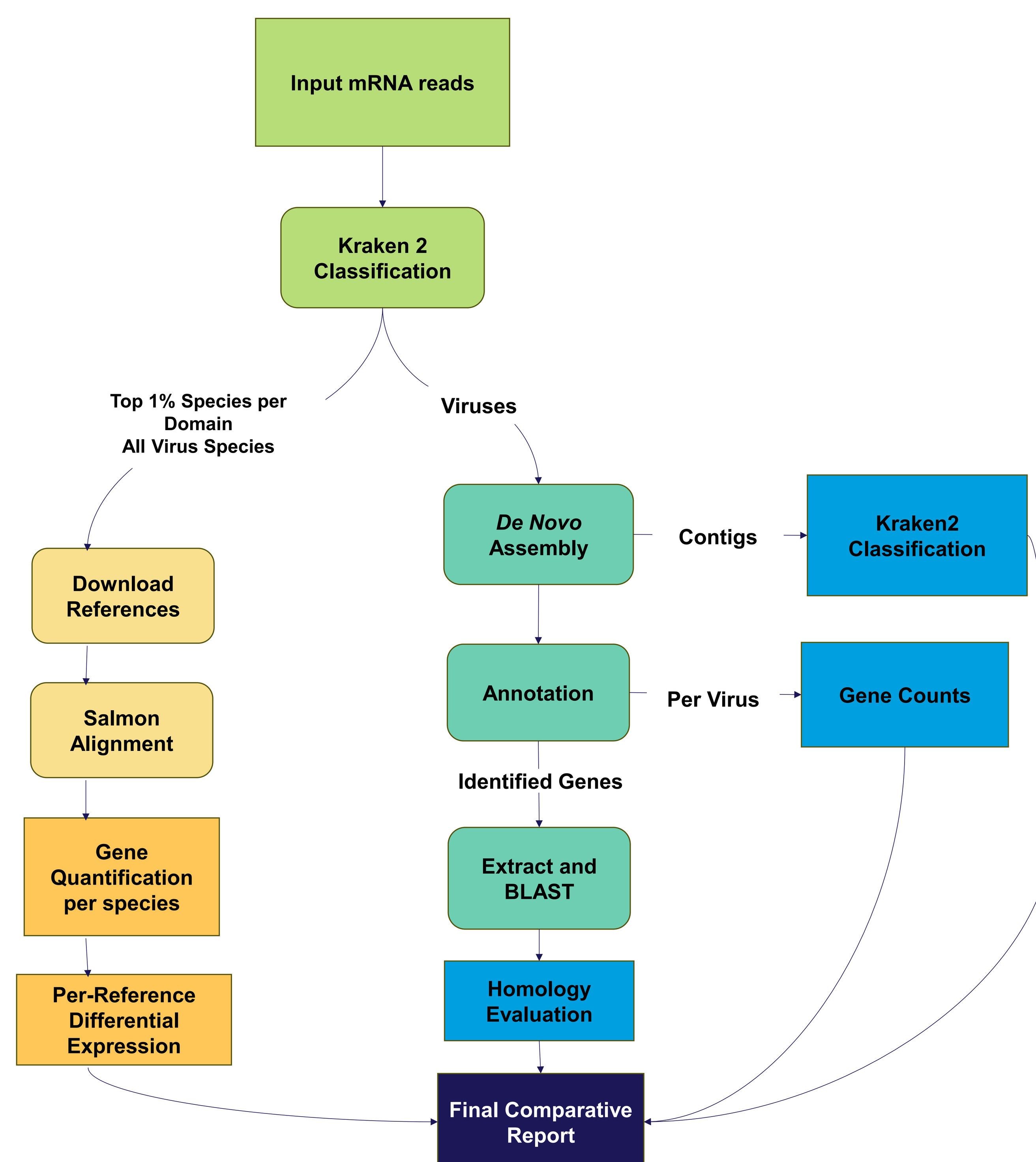

Figure 1. Overview of bioinformatics workflow. Reference selection is based on identifying non-host organisms and downloading sequence and annotation information for organisms that represent at least 1% of their respective domain-level classification. Differential expression was then performed on a reference by reference basis.

RESULTS

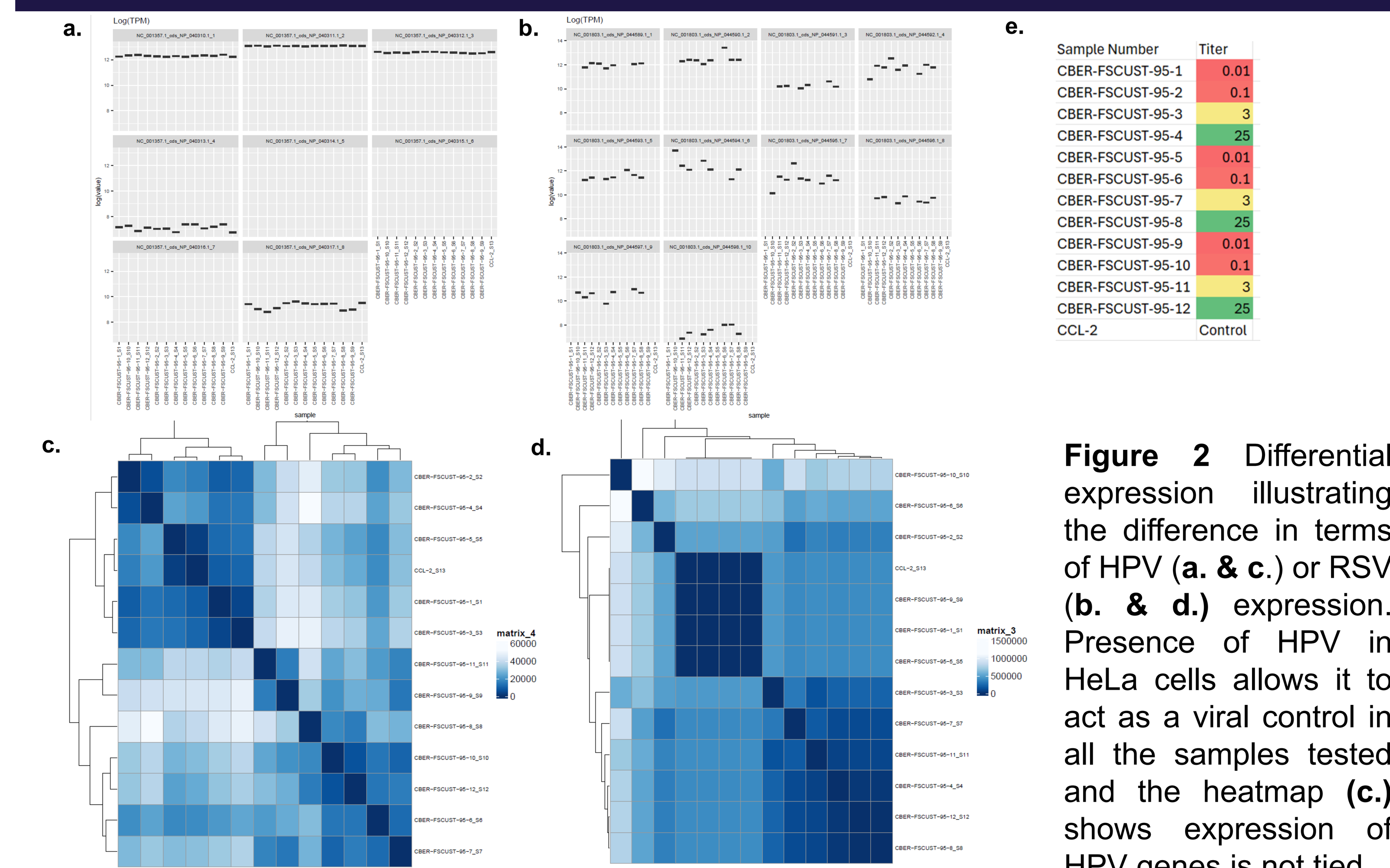

Figure 2 Differential expression illustrating the difference in terms of HPV (**a. & c.**) or RSV (**b. & d.**) expression. Presence of HPV in HeLa cells allows it to act as a viral control in all the samples tested and the heatmap **(c.)** shows expression of HPV genes is not tied to the levels of RSV added to each sample compared to the heatmap for RSV genes (**d.**) which show clustering based on titer of the virus added. This data also shows that the negative control sample clusters most closely with the samples that have the lowest titer of RSV spiking (0.01gc/cell). The box plots show how the genes for HPV **(a.)** is much more consistent among titers of spiked sample compared to the RSV spiked samples **(b.)**. Scale **(e.)** is a simple legend showing the levels of spiked sample with accompanying titers.

CONCLUSION

- We are developing a comprehensive pipeline for the detection of adventitious agents using NGS technology, which has the potential to replace traditional in vivo detection methods. This innovative approach aims to address the limitations of current practices, such as the ethical concerns and high costs associated with live animal models, and the restricted scope of PCR-based techniques.
- Our research involves spike studies with a diverse array of viruses, representing different viral groups and genome types. Through these studies, we have established the detection limits for specific viruses of interest. These results can be extrapolated to predict the detection capabilities for other similar viruses, enhancing the robustness of our method.
- A key component of our approach is the analysis of differential gene expression in spiked cell samples. By comparing the gene content of these samples to that of a control, we can identify significant changes that may indicate the presence of adventitious agents. This method allows us to focus on specific genes of interest, thereby reducing the likelihood of false positives and improving the accuracy of our detection system.
- The standardization of our methods is essential to ensure that our virus detection system operates as intended. This involves rigorous testing to confirm that the system can differentially and selectively identify potential adventitious agents. By establishing a reliable and precise detection pipeline, we aim to significantly improve the safety and efficacy of biological product manufacturing.

Acknowledgments: We would like to thank AFS leadership, especially Dr. Joseph Leonelli and Rebecca Bradford, for their continued support of this research , as well as the ATCC Internal Research and Development program.

© ATCC 2024. The ATCC trademark, trade name, any and all ATCC catalog numbers listed in this publication are trademarks of the American Type Culture Collection unless indicated otherwise.

AFS 10801 University Boulevard, Manassas, VA 20110-2209 **Phone:** 800.359.7370 **Email:** federalsolutions@atcc.org **Web:** https://www.atcc.org/federal-solutions

. © 2023 American Type Culture Collection. The ATCC trademark and trade name, and any other trademarks listed in this publication are trademarks owned by the American Type Culture Collection unless indicated otherwise.