LB-9254

Use of NGS-Based Minor Variant Analysis to Monitor Genomic Stability in Clinical Viral Isolates Intended for Challenge Studies

Manasi Tamhankar, Ann Wasko, Nikhita Puthuveetil, Amy Reese, Angela Robertson, and **Sujatha Rashid**

BEI Resources, ATCC, Manassas, Virginia, USA.

INTRODUCTION

The availability of viral stocks that are genetically and functionally similar to the original clinical isolates is crucial for developing effective medical countermeasures (MCM). A critical, and unavoidable, aspect of generating these isolates involves using immortalized tissue culture cells for virus production. However, multiple passages of viruses can induce tissue-culture adaptation, leading to phenotypic alterations in the viral isolates. Such adaptations can compromise the evaluation of MCM efficacy, underscoring the necessity of monitoring the genetic stability of virus isolates derived from clinical samples.

The optimization of virus growth conditions in cell culture typically focuses on obtaining high titers and rarely on genomic fidelity of the resulting viral strain. Monitoring minor variants present in viral populations provides crucial data for selecting the optimal cell line that supports robust viral replication while preserving genomic fidelity.

Here, we optimized virus production (human coronavirus hCoV-229E and Oropouche virus) by using next-generation sequencing (NGS)-based minor variant analysis to monitor the genomic stability of clinical isolates during growth in tissue culture.

METHODOLOGY

- For each passage, viral RNA was extracted from the infected cell culture supernatant using QIAamp viral RNA mini kit (QIAGEN).
- Libraries were prepared using the DNA prep kit, indexed with DNA/RNA UD indexes, and whole-genome paired-end sequencing was performed on an Illumina MiSeq to a minimum 100× average coverage.
- Base-calling and adapter trimming were done using Illumina onboard software, followed by additional trimming and quality filtering with fastp and FastQC.
- Trimmed reads were mapped to the appropriate reference genome using bwa mem.
- Viterbi realignment was applied with LoFreq to correct mapping errors.
- Initial variant calling was done using bcftools and variants with PHRED ≥30 were retained. Base quality score recalibration (BQSR) was done with Picard and GATK. Coverage was assessed using Qualimap. Supplemental alignments were removed with SAMTools to eliminate chimeric mappings.
- Variant-calling was repeated using LoFreq on the BQSR-corrected BAM, filtering for ≥5% allele frequency and ≥10× coverage.
- A consensus sequence was generated by identifying low coverage regions with bedtools and applying a custom Python script using reference, LoFreq VCF, and coverage data.
- Variants were compared to the GenBank Flat File (GBK) of the reference genome to localize changes and calculate amino acid substitutions using a custom Python script.

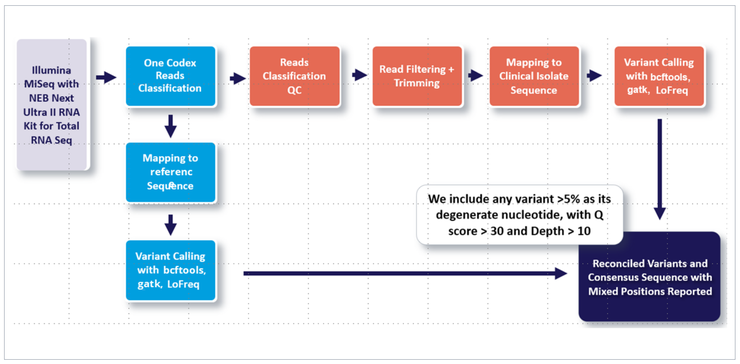

Figure 1: NGS and Minor Variant Analysis for RNA viruses

RESULTS – hCoV-229E

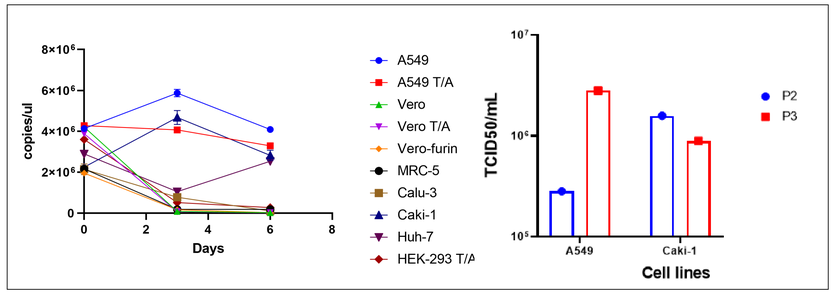

Figure 2. Screening and Replication of hCoV-229E in Immortalized Cell Lines. A passage 3 hCoV-229E UNC/2/2022 isolate grown in human airway epithelial air-liquid interface (HAE-ALI) cultures was used for all experiments. **(A)** Ten cell lines were screened for hCoV-229E replication at 37°C using qPCR and dsRNA immunostaining. **(B)** Replication kinetics in A549 and Caki-1 cells following serial passage. Infectivity of passages 2 and 3 was quantified by $TCID_{50}$ assays

- Four cell lines, A549 (CCL-185™), A549-TMPRSS2/ACE2, Caki-1 (HTB-46™), and Calu-3 (HTB-55™), demonstrated robust viral replication.
- Virus propagation in A549 and Caki-1 cells resulted in high viral titers (Figure 2A).
- In A549 cells, a ten-fold increase in viral titer was observed between passage 2 and passage 3, indicating enhanced replication efficiency upon serial passage (Figure 2B).
- In contrast, viral titers in Caki-1 cells remained stable across passages.

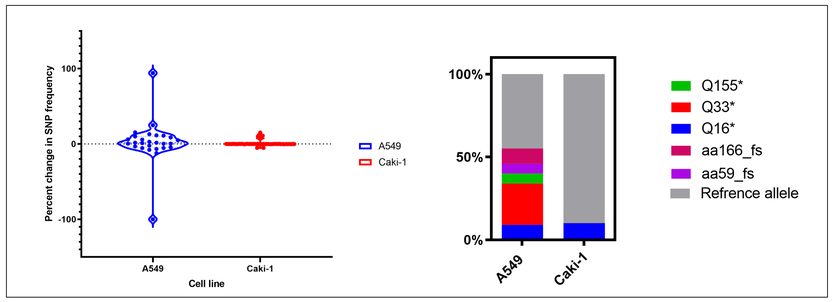

Figure 3. Variant Analysis and ORF4 Integrity After Serial Passage. (A) Percent change in the frequency of SNPs seen across the hCoV-229E genome following 3 passages in A549 or Caki-1 cells. **(B)** Variant frequency of truncating SNPs in the ORF4 gene in Caki-1-passaged virus showed minimal changes.

- Across all cell lines tested, the hCoV-229E virus accumulated a range of single-nucleotide polymorphisms (SNPs) distributed across the genome. The majority of these SNPs were low frequency (<15%) and would not be apparent in the consensus sequence (Figure 3A).
- The hCoV-229E UNC isolate initially possessed a contiguous ORF4, which remained intact through two passages in HAE-ALI cultures. However, by passage 3 in immortalized cell lines, ORF4 exhibited extensive mutational changes whose frequency and nature varied by cell line.
- In A549 cells, passage 3 virus harbored multiple mutations within ORF4, including those resulting in premature stop codons and frameshift events, indicative of strong selective pressure leading to truncation of the encoded protein.
- In contrast, virus passaged in Caki-1 cells accumulated only a single low-frequency SNP introducing a stop codon, suggesting a comparatively reduced mutational burden (Figure 3B).

RESULTS - OROV

- OROV propagated in Vero cells exhibited notable changes in SNP frequencies between passages (Figure 4), including variants disrupting the start codon. In contrast, OROV grown in A549 cells showed no significant changes, indicating greater genomic stability.
- Growth in Vero cells resulted in multiple changes across all viral segments, including amino acid substitutions in the M segment encoding the glycoprotein envelope and variants affecting the start codon (methionine to threonine). No such changes were observed in A549 cells (Tables 1A and 1B).

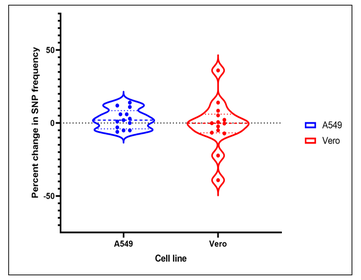

Figure 4: Percent change in the frequency of SNPs across the OROV genome following passage in A549 and Vero cells. OROV was serially passaged twice in A549 (CCL-185™) and Vero (CCL-81™) cells.

Vero (3.8E+06 $TCID_{50}$/mL)

Segment	Position variant from P0	P3	P4	Change	AA mutation
S	836_837insA	12%	12%	0.60%	-- (3'UTR)
S	947A>G	39%	0%	-39%	-- (3'UTR)
M	32T>C	0%	5%	5%	M1T
M	1656A>G	7%	9%	2%	--
M	1860C>T	16%	14%	-2%	--
M	1872A>C	28%	5%	-22%	K614N
M	2272A>G	16%	52%	36%	K748N
M	2562G>A	5%	13%	8%	--
L	37T>C	0%	14%	14%	M1T
L	222T>C	100%	99%	-1%	F63L
L	1781C>T	70%	69%	-1%	--
L	3294_3295insA	5%	0%	-5%	frameshift
L	4698A>G	6%	0%	-6%	I1554G
L	5027A>G	19%	11%	-7%	--

A549 (8.9E+06 $TCID_{50}$/mL)

Segment	Position variant from P0	P3	P4	Change	AA mutation
S	836_837insA	13%	26%	12%	-- (3'UTR)
S	837del	0%	5%	6%	-- (3'UTR)
S	947A>G	0%	10%	11%	-- (3'UTR)
M	1685C>T	0%	13%	14%	A552V
M	1860C>T	20%	22%	3%	--
M	1872A>C	5%	0%	-5%	K614N
M	2562G>A	20%	21%	2%	--
M	4279_4280insA	5%	0%	-6%	frameshift
L	17del	5%	0%	-5%	-- (5' UTR)
L	222T>C	100%	100%	0%	F63L
L	1781C>T	65%	62%	-3%	--
L	2877_2878insA	0%	5%	6%	frameshift
L	5027A>G	23%	24%	1%	--

Tables 1A and 1B: Minor variants detected by NGS across passages 3 and 4 relative to the original sequence (P0) of the OROV clinical viral isolate.

SUMMARY

- We successfully applied our custom NGS-based minor variant analysis pipeline to evaluate the impact of tissue culture propagation on recent clinical isolates of Oropouche virus and human coronavirus hCoV-229E.
- For coronaviruses, potential tissue-culture-associated genetic changes were identified in as little as two passages and allowed for the detection SNPs before they manifested in consensus sequences and induced phenotypic changes.
- In general, the detection of minor genetic variants can serve as an early indicator of tissue-culture adaptation, underscoring the utility of such NGS-based analysis for monitoring genomic stability in clinical viral isolates across passages.
- These results highlight the importance of monitoring genomic fidelity alongside virus titer while selecting a cell line for virus propagation.
- It is recommended that viral isolates that are intended for developing MCMs should be maintained at lower passages. Moreover, monitoring of the emergence of minor variants during the production of viral isolates from clinical samples will maximize their suitability for the development of therapeutics.

ACKNOWLEDGEMENTS

This work was funded through BEI Resources (75N93024C00001) by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services.