Vaccine 37 (2019) 1001–1005

Contents lists available at ScienceDirect

Vaccine

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

Towards dynamic monitoring of cell cultures using high throughput sequencing

Shasta D. McClenahan [1], Philip R. Krause *

Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD USA

ARTICLE INFO

Article history:
Received 5 July 2018
Received in revised form 4 December 2018
Accepted 9 December 2018
Available online 11 January 2019

Keywords:
Adventitious agent
Virus
Degenerate PCR
High throughput sequencing (HTS)

ABSTRACT

We used a combination of DOP-PCR with high throughput sequencing (HTS) to study infected cell cultures over time to assess the feasibility of using this technique to provide a read-out other than cytopathic effect in cell culture infectivity assays. Because DOP-PCR primers feature a short constant sequence at their 3′ terminus, the procedure yields a reproducible representational library of products from a given PCR template, including viral nucleic acids. Using SV40- and MVM-infected cultures harvested at different times, we show that the number of viral matches among DOP-PCR products parallels the quantity of virus as shown by real-time PCR, and further show that HTS analysis of specific DOP-PCR products that increase in quantity over time could be used to identify the infecting virus with a sensitivity similar to that of typical cell-culture assays that rely on cytopathic effect.

1. Introduction

Assurance that adventitious agents are not present in the manufacturing processes for biological materials including vaccines and therapeutic products is an important element of product quality control. While adventitious agents, potentially introduced via source materials, environmental factors, or cell substrates, are rare, they can have important public health consequences, including the potential to interrupt product supply.

Examples of adventitious agents in biological products include the identification of porcine circovirus (PCV) by high-throughput sequencing in rotavirus vaccines in 2010 [1,2]. Minute virus of mice (MVM) was reported as an adventitious virus in Chinese hamster ovary (CHO) cells used for production of biotherapeutic products [3]. Simian virus type 40 (SV40) was also detected in poliovirus vaccines in 1960 as a result of its presence in primary monkey kidney cells used to produce the vaccine [4].

Rapid advances in high-throughput sequencing (HTS) raise the possibility that HTS could be used in quality assurance for adventitious agents in biological products. Potential concerns about the use of HTS include the risk of false positives (with potential impact on product supply), the need for time-consuming follow-up of questionable signals, or detection of unknown sequences (which could potentially lead to repeated bioinformatics analyses as new sequences are added to the databases). Thus, to date, while HTS has been used in qualifying cell substrates on a one-time basis [5,6] and for post-hoc investigation of potential viral contamination of cell cultures, HTS-based techniques have not found widespread acceptance in a good manufacturing process (GMP) stream.

Currently-used adventitious agent detection methods include cell-culture assays, polymerase chain reaction (PCR) assays, electron microscopy, biochemical tests (e.g., for reverse transcriptase), and animal testing [7]. Because of their ability to detect viable virus, including those that can grow in the cell substrate, cell-culture assays have been at the core of adventitious agent detection strategies, recognizing that some viruses (e.g., PCV) may not be cytopathic and that the read-out of cytopathic effect (CPE) often used for these assays is not always perfectly objective.

The combination of cell-culture assays with HTS could potentially overcome some of the limitations of each of these assays, based on the presumption that sequences that increase over time in a cell-culture assay are most likely to represent adventitious agents. A “time zero” sample could serve as a negative control for the HTS assays, eliminating the need to consider sequences that do not increase over time. Cell-culture assays combined with HTS would serve as a built-in infectivity assay for sequences observed in HTS, permitting focus on sequences that truly have the potential to represent infectious virus. While HTS has been used to evaluate cells directly for the presence of adventitious agents, this is the first

---

* Corresponding author at: WO71 RM 3234, HFM-457, 10903 New Hampshire Ave, Silver Spring, MD 20993-0002, USA.

E-mail address: Philip.krause@fda.hhs.gov (P.R. Krause).

[1] Present address: National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 1315 East West Highway, SSMC3 Room 13319, Silver Spring, MD, USA.

https://doi.org/10.1016/j.vaccine.2018.12.019
0264-410X/Published by Elsevier Ltd.

report of using this technique to use this approach to non-specifically monitor cells over time for the presence of sequences that could potentially represent infecting virus.

In the present study, we performed proof-of-concept experiments to determine whether HTS can be sufficiently quantitative for potential use in detecting a replicating virus and identifying specific sequences that increase over time in cell-culture assays. We amplified nucleic acids obtained from particle-enriched (using ultracentrifugation and nuclease digestion) cell-culture infectivity assays with degenerate-oligonucleotide primer (DOP) PCR and subjected the resulting DOP-PCR products to HTS using the MiSeq platform. The use of DOP-PCR, which is biased towards specific sequences that match a short 3′ sequence on each primer, provides a reproducible read-out of specific sequences representing all sequences present in a sample, simplifying the bioinformatics analysis to identify sequences that increase over time and comparison with sequences in GenBank. We previously showed that DOP-PCR could be used to sensitively detect viruses spiked into or growing in cell culture [5,8], and this technique has also been used to identify unknown viruses in clinical samples [9,10] and cell culture [11]. DOP-PCR is adaptable to HTS, as use of multiplexing barcodes and addition of Illumina sequencing primers are facilitated by PCR [5]. We also explored approaches to normalizing results to account for potential sample-to-sample differences in nucleic acid extraction, the DOP-PCR, and the sequencing itself.

2. Materials and methods

2.1. Infectivity assays

Chinese hamster ovary cells (CHO, ATCC CCL-61) were infected with minute virus of mice (MVM; ATCC VR-1346) and African green monkey kidney cells (Vero, ATCC CCL-81) were infected with simian virus type 40 (SV40, a gift of Keith Peden). The cells were seeded in Modified Essential Medium (MEM, Gibco) supplemented with 2% fetal bovine serum (FBS) in the 6-well plates at a low cell density (1×10^5 cells/well) to allow for continual growth for up to five days of viral infection. The cells were allowed to attach for 2 h prior to infection.

The model adventitious viruses were serially diluted ten-fold in MEM and stored on ice until used. The cell monolayers were rinsed twice with MEM, and 1 mL of each virus dilution was added to each well, along with a negative control MEM only well. The plates were incubated for 2 h at 37 °C with gentle rocking every 15 min. Following the 2-h infection, the inoculum was removed, and the cells were washed twice with MEM. Fresh MEM with 2% FBS was added, and the plates were returned to the 37 °C incubator for the duration of the experiment. An entire well was collected at each time point starting at time 0, and then daily for two days following MVM infection and five days following SV40 infection. The cells were scraped in the well and the cells and supernatant were harvested into 2 mL microfuge tubes. The cells were pelleted and all but 500 μL of supernatant was transferred to a new collection tube. The cell pellets and supernatants were stored at −70 °C until the experiment was completed.

2.2. Sample preparation and high throughput sequencing

Nucleic acids were prepared as previously described [5,8,12] from freeze-thawed cells or supernatant using the All Prep Kit (Qiagen, Valencia, CA). Cell pellets were frozen and thawed three times and then passed through a QiaShredder column (Qiagen). Samples were relatively enriched for virus particles by nuclease digestion and ultracentrifugation as described previously [2,9,12]. Briefly, samples were digested with DNase I (Sigma-Aldrich, St. Louis, MO) for 30 min at 37 °C. The samples were then centrifuged through a 1 M NaCl, 10 mM Tris-HCl, pH 7.4, cushion at 30,000 × g at 4 °C for 90 min. After removal of the supernatant, nucleic acids were extracted from the pellet with the All Prep kit according to the manufacturer's instructions.

T2 or T4 bacteriophage (ATCC 11303-B2 or 35060-B1, respectively) were spiked into the samples to normalize for potential variation in efficiency of capsid preparation, DOP-PCR, and sequencing of virus-sized genomes. Lyophilized phage were resuspended in sterile water according to ATCC instructions and diluted to 100–1000 PFU/mL in MEM. One hundred μL of each sample was spiked into the cell pellets or supernatants prior to DNA or RNA extraction.

Reverse-transcribed RNA (using random hexamers) or DNA from the particle enrichment procedure was subjected to DOP-PCR using a single PCR primer containing a specific 3′ sequence (to allow DOP-PCR), as well as a 5′ sequence designed to permit addition of the Illumina sequencing primers in a subsequent PCR. Final libraries were prepared for HTS using a primer with an index (or barcode) for HTS, followed by an additional overlapping PCR reaction to add the full Illumina sequencing primer sequence containing the dual indexes (D501-508 and D701-712), as previously described [8]. Primers used are described in Table 1.

HTS was performed on an Illumina MiSeq instrument used in-house with 2 × 300 nucleotide read lengths and pooled samples using dual indexes.

2.3. Sequence analysis

Sequence analysis was automated using the National Institutes of Health Biowulf 2 supercomputing cluster. Raw sequences were trimmed of primer sequences using fastx_clipper and fastx_trimmer, followed by conversion to fasta format using fastq_to_fasta. Using in-house python scripts, identical sequences in each run were consolidated and sorted by frequency, and specific sequences that increased over time were identified by comparing the frequency of each sequence in the results from two different time

Table 1
PCR primers used for these experiments. Index sequences were added to some samples for multiplexing.

Primer Name	Application	Sequence	Sense
DOP	Nonspecific amplification	CCGACTCGAGINNNNNNTGTGG	+
IL-DOP	Nonspecific Illumina amplification	GCTCTTCCGATCTINNNNNNTGTGG	+
Illumina P5 Adapter	Overlap PCR for Illumina library	AATGATACGGCGACCACCGAGATCTACAC-*Index*-ACACTCTTTCCCTACACGACGCTCTTCCGATCT	+
Illumina P7 Adapter	Overlap PCR for Illumina library	CAAGCAGAAGACGGCATACGAGATC-*Index*-GACTGGAGTTCAGACGTGTGCTCTTCCGATC	–
MVM F	Specific PCR Forward primer	AGTTTGCCATGCTATTTGC	+
MVM R	Specific PCR Reverse primer	ACTGGTTTACTTGCTGTCC	–
MVM Probe	qPCR Probe	FAM-ATTTCTTTTGCCTCCTTGTCTGTTT –TAMRA	+
SV40 F	Specific PCR Forward primer	GTGGAATGCCTTTAATGAGGAAA	+
SV40 R	Specific PCR Reverse primer	TTGAGAGTCAGCAGTAGCCTCATC	–
SV40 Probe	qPCR Probe	FAM-CCTGTTTTGCTCAGAAGAAATGCCATCTAGTG- TAMRA	+

points, calculating the fold-increase over time, and retaining sequences that showed statistically significant increases (at a p value of 0.01 by Chi Square) after normalization to phage sequences detected in the same results. Sequences observed to increase over time were characterized by BLASTN search both for sequences matching the starting virus reference genome or sequences in the GenBank nt database.

2.4. Real-time quantitative PCR

The viral DNA or RNA from infected cells or supernatant was quantified by Taqman (Applied Biosystems) real-time PCR (qPCR). Ten-fold serial dilutions of plasmid containing viral sequence were used as standards to quantify the virus from the infectivity experiments. A total of 5 or 10 μL of RNA or DNA from the DE or CP material was used in each PCR reaction in triplicate. The qPCR mastermix included 2x mastermix (Applied Biosystems), and primers and probes specific for each virus target (Table 1).

3. Results

Cell cultures were harvested at day 0 (directly after virus adsorption) after infection with model adventitious viruses SV40 or MVM, and at different numbers of days after inoculation of the cultures. Timing of any CPE was observed. To aid in normalization, harvested samples were spiked with DNA phages T2 or T4. Samples were enriched for encapsidated nucleic acids using nuclease digestion and ultracentrifugation. Samples were tested at each time point using quantitative real-time PCR (by Taqman) to quantify the infecting virus and by DOP-PCR followed by HTS. To reduce DOP-PCR/HTS assay variability, we included all related samples in the same sequencing run using multiplexing barcodes, and also normalized results to those obtained with spiked phage. The quantities of sequences matching the infecting virus were counted, and sequences detected at each time point were also compared with quantities of those same sequences that were detected at time 0 to determine if HTS-based techniques for virus detection could potentially be used to identify an adventitious virus growing in cell culture.

Ten-fold serial dilutions of SV40 containing 1×10^1 – 1×10^6 PFU/mL were used to infect Vero cells in 6-well plates. The CPE consistent with SV40 infection became apparent at each dilution by 2–3 days after inoculation (Fig. 1A). T4 phage was spiked into each sample prior to extraction. SV40 DNA was quantified by qPCR at each time point (Fig. 1B), with at least 1 log increases in DNA observed by 2 days after inoculation at each dilution. SV40 DNA quantities continued to increase over time, though generally at lower relative rates between days 3 and 5.

Samples from cultures inoculated with SV40 at 1×10^2 (Fig. 2A) and 1×10^3 PFU/mL (Fig. 2B) were subjected to HTS at each time point. After normalization to the number of T4 phage sequences, the quantity of sequences in the HTS run with high homology (E value < 10^{-50}) to the reference SV40 sequence ("MiSeq") paralleled the qPCR results at each time point. The quantity of sequences identified by restricting the analysis just to the DOP-PCR product sequences that increased significantly over time ("Increasers"), as well as the quantity associated with the individual DOP-PCR product sequence that increased the most over the period from time 0 to each subsequent time point ("Greatest Increaser"), also parallels the qPCR results. The most abundant "increasers" represented SV40 sequences at both inocula and all time points, except for on day 2 at the 10^2 PFU inoculum, at which the most abundant increasing sequence was a mitochondrial sequence, and the second-most abundant sequence represented SV40. Thus, restricting the HTS analysis to significantly increasing sequences was able to detect the SV40 infection by day 2, at least as quickly as was possible by CPE or by other means.

Ten-fold serial dilutions of MVM containing approximately 1×10^0 –1×10^3 TCID$_{50}$/mL were used to infect CHO cells. At the

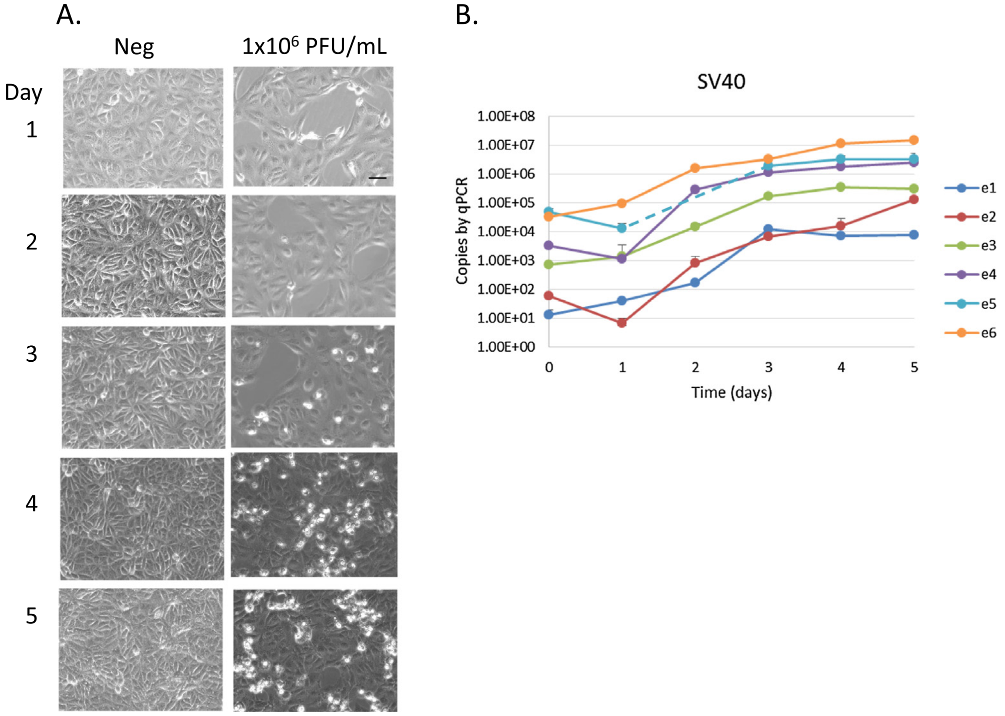

Fig. 1. Evaluation of Vero cells infected with SV40. A) Cells were infected at day 0 at 1×10^6 PFU/mL and observed daily for 5 days. Neg: Uninfected control. Scale bar: 50 μM. B) Cells infected at different inocula were evaluated for SV40 sequences by Taqman PCR daily for 5 days after inoculation.

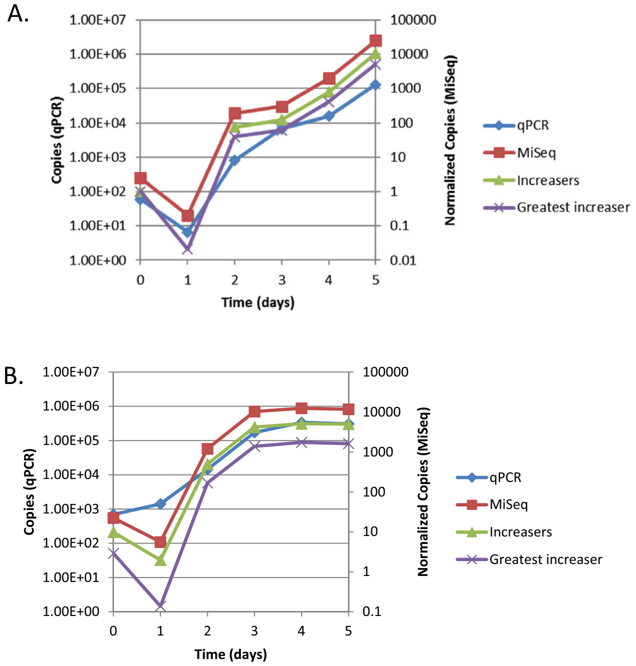

Fig. 2. Comparison of qPCR and MiSeq results for SV40. qPCR results from Fig. 1B (shown on left axis) are compared with MiSeq (shown on right axis) results for infecting inocula of A) 10^2 PFU/mL and B) 10^3 PFU/mL. MiSeq results are normalized to spiked T4 phage, and presented as total SV40 sequences (MiSeq), statistically significant increasing sequences that were attributable to SV40 (increasers), and the single most abundant significantly increasing sequence attributable to SV40 (greatest increaser). Significantly increasing sequences were first apparent at 2 days after infection.

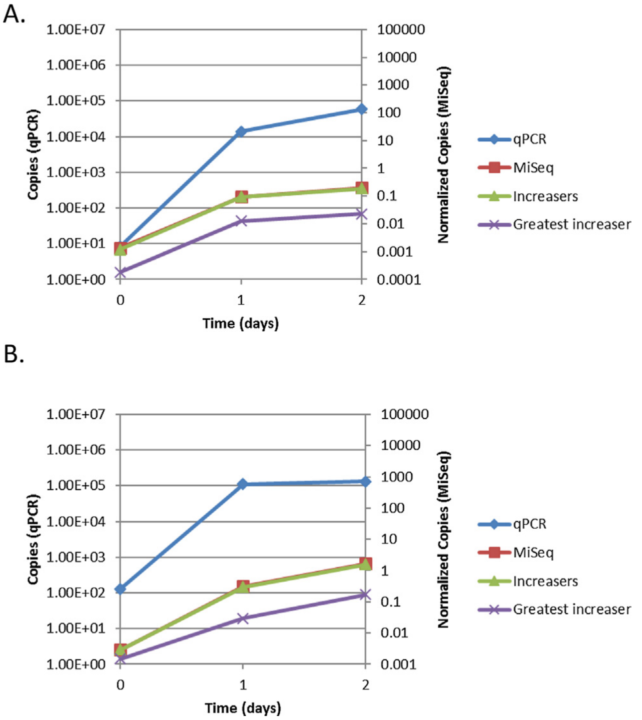

Fig. 4. Comparison of qPCR and MiSeq results for MVM. qPCR results from Fig. 3B (left axis) are compared with MiSeq results (right axis) for C) 10^2 PFU/mL and D) 10^3 PFU/mL. MiSeq results are normalized to spiked T2 phage, and presented as total MVM sequences (MiSeq), statistically significant increasing sequences that were attributable to MVM (increasers), and the single most abundant significantly increasing sequence attributable to MVM (greatest increaser). The first significant increase was observed at 1 day post-inoculation, and the quantity of those specific sequences at Day 0 is also graphed.

higher inocula, the cells showed evidence of infection, with increased cellular refractility and decreased growth relative to control, by one day after inoculation (Fig. 3A). T2 bacteriophage was spiked into the samples prior to extraction of DNA. By qPCR, MVM DNA increased by over 2 logs from the inoculum at day 1, with continuing increases by day 2 for the cultures inoculated at 10^1 and 10^2 TCID$_{50}$/mL (Fig. 3B).

Samples from cultures inoculated with 10^2 (Fig. 4A) and 10^3 (Fig. 4B) TCID$_{50}$/mL of MVM were subjected to HTS. After normalization to the number of spiked T2 phage sequences, the number of MVM sequences detected by HTS paralleled the increases observed by qPCR. Of these sequences, the quantity identified by restricting the analysis just to sequences that increased significantly over time is also graphed, as well as the quantity associated with the individual sequence that increased the most over the period from time 0 to each subsequent time point. The most abundant significantly increasing sequences all represented MVM sequences at all time points including at one day after inoculation. Thus, restricting the HTS analysis to significantly increasing sequences was able to detect the MVM infection by day 1, similar to observation of cytopathic effect in culture.

4. Discussion

We evaluated growth over time in cell culture of SV40 and MVM by evaluation of CPE, qPCR, and HTS. We found that, after

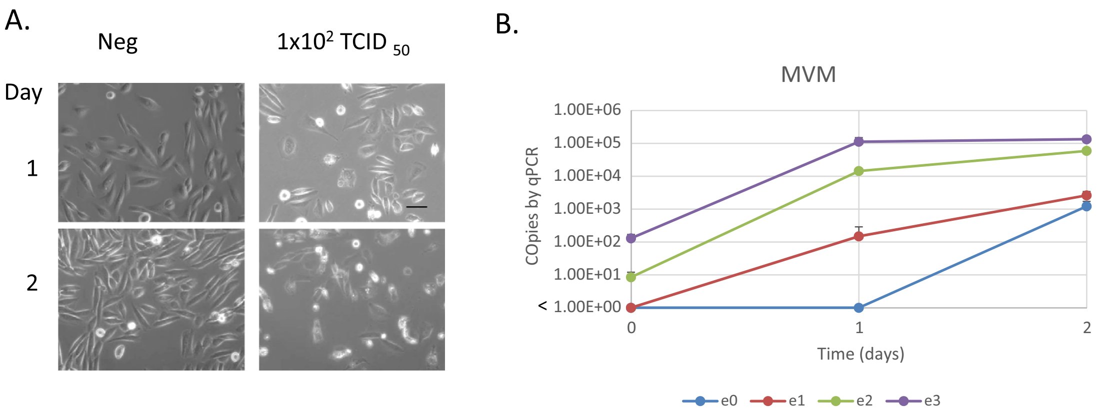

Fig. 3. Evaluation of CHO cells infected with MVM. A) Cells were infected at day 0 at 1 × 10^2 TCID$_{50}$ and observed daily for 2 days. Neg: Uninfected control. Scale bar: 50 μM. B) Cells infected at different inocula were evaluated for MVM sequences by Taqman PCR daily for 5 days after inoculation.

normalization to sequences of spiked phage, HTS results paralleled those observed by qPCR in infected cell culture of SV40 and MVM. We further showed that comparison of sequence numbers at late time points vs post-inoculation time points permitted identification of the infecting virus for all of these viruses, as quickly or sooner than the emergence of CPE in the cultures.

We paid particular attention to enhancing the consistency of the HTS results, both by experimental design and by normalizing results to those obtained by spiking phage into the samples prior to extraction for DOP-PCR and HTS. We considered several approaches to minimizing artifacts that could arise due to differences in sample preparation, DOP-PCR, or HTS. We improved the consistency of HTS by using multiplexing adapters and testing all samples from a single infection within a single run.

We considered several approaches to normalizing the results to further improve consistency, including normalizing to total sequences obtained in each infection at each time point, normalizing only to cellular sequences obtained at each time point, or normalizing to a spiked control. We noted that normalizing to total sequences would account for differences in sequencing but not in DOP-PCR or in nucleic acid extractions. Normalizing to cellular sequences identified in the sequencing output could be performed by identifying cellular sequences by taxonomic identification or a string search within the BLAST results, and would have the advantage of assuring that any identified viral sequences "outgrew" cellular sequences, which are not expected to increase over time, but this also would not account for potential differences among samples in the particle enrichment by CP. We chose to use a phage spike, which, because phages are also purified by typical viral particle enrichment techniques, can account for more sources of variability, including sample-to-sample variability from the particle enrichment, the DOP-PCR amplification, and sequencing variability. We used DNA bacteriophage for this spike, which allowed simple normalization of results from DNA viruses. We did not investigate whether or not an RNA phage (e.g., MS2) could be used to normalize detection of RNA virus sequences, but hypothesize that this approach would also be useful for RNA viruses.

We noted that many of the cellular sequences observed were mitochondrial, suggesting that mitochondria may co-purify with viruses in the particle enrichment, but not be adequately normalized by phage spike. There also were cellular sequences that appeared to increase without a clear explanation (although at rates lower than the viral sequences)—possibly as a result of experimental variability. If cellular sequences are observed using this technique, because the cells that are used for infectivity assays are known and are banked, libraries of cellular sequences that are susceptible to this type of variability could be identified in advance on non-product-related samples. Such sequences can also serve as an internal control since real viruses should increase over time both relative to phage and to cellular sequences.

We note several limitations of this approach. Sequences that do not increase over time are unlikely to be viruses, but if the culture is already highly infected at time 0 (as was the case when porcine circovirus 1 was identified in 2010 as an adventitious agent of rotavirus vaccines), further increases in these sequences may not be observed. RNA viruses may be more difficult to detect if cultures are examined only at time 0 and one other time point, since RNA is less stable than DNA and lytic infection may release enzymes that destroy viral RNA over time. However, lytic infection is normally detectable by CPE.

For SV40, we found significant increases in viral sequences around the time CPE became apparent, while for MVM, significant increases in viral sequences were apparent the day after inoculation, implying sensitivity of this approach comparable to that of cell culture. While MVM may grow with different kinetics in serum-free medium (as compared with the serum-containing medium used in our experiments), it seems unlikely that this would change the relative sensitivity of cell culture CPE read-out vs. identification of sequences that increase relative to time 0. In theory, the sensitivity of this approach could be further increased by increasing the number of sequences obtained from each sample.

We conclude that, in principle, HTS could be used in conjunction with existing cell-culture assays to improve the utility of both assays. With appropriate normalization, HTS appears to be sufficiently quantitative to identify infecting viruses. These findings suggest the possibility that this type of approach ultimately could be used to directly monitor cultures for sequences that increase during actual production of biological products.

Conflict of interest statement

The authors have no financial or personal relationships that influenced this research. Funding was provided by FDA.

Acknowledgements

The authors would like to thank Drs. Haru Murata and Keith Peden (FDA) for providing SV40 virus for these experiments, and Drs. Steven Rubin and Keith Peden for critical reading of the manuscript. This research was supported by the FDA/CBER intramural research program.

References

[1] Victoria JG et al. Viral nucleic acids in live-attenuated vaccines: detection of minority variants and an adventitious virus. J Virol 2010;84:6033–40.

[2] McClenahan SD, Krause PR, Uhlenhaut C. Molecular and infectivity studies of porcine circovirus in vaccines. Vaccine 2011;29:4745–53.

[3] Garnick RL. Experience with viral contamination in cell culture. Dev Biol Stand 1996;88:49–56.

[4] Sweet BH, Hilleman MR. The vacuolating virus, S.V. 40. Proc Soc Exp Biol Med 1960;105:420–7.

[5] McClenahan SD, Uhlenhaut C, Krause PR. Evaluation of cells and biological reagents for adventitious agents using degenerate primer PCR and massively parallel sequencing. Vaccine 2014;32:7115–21.

[6] Onions D et al. Ensuring the safety of vaccine cell substrates by massively parallel sequencing of the transcriptome. Vaccine 2011;29:7117–21.

[7] US Food and Drug Administration, Guidance for Industry: Characterization and Qualification of Cell Substrates and Other Biological Materials Used in the Production of Viral Vaccines for Infectious Disease Indications. https://www.fda.gov/downloads/biologicsbloodvaccines/guidancecomplianceregulatoryinformation/guidances/vaccines/ucm202439.pdf (2010).

[8] McClenahan S, Uhlenhaut C, Krause P. Optimization of virus detection in cells using massively parallel sequencing. Biologicals 2014;42:34–41.

[9] Uhlenhaut C et al. Use of a novel virus detection assay to identify coronavirus HKU1 in the lungs of a hematopoietic stem cell transplant recipient with fatal pneumonia. Transpl Infect Dis 2012;14:79–85.

[10] Uhlenhaut C, Cohen JI, Fedorko D, Nanda S, Krause PR. Use of a universal virus detection assay to identify human metapneumovirus in a hematopoietic stem cell transplant recipient with pneumonia of unknown origin. J Clin Virol 2009;44:337–9.

[11] McClenahan SD et al. Discovery of a bovine enterovirus in alpaca. PLoS ONE 2013;8:e68777.

[12] Nanda S et al. Universal virus detection by degenerate-oligonucleotide primed polymerase chain reaction of purified viral nucleic acids. J Virol Methods 2008;152:18–24.