Biologicals 42 (2014) 34–41

Contents lists available at ScienceDirect

Biologicals

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

# Optimization of virus detection in cells using massively parallel sequencing

Shasta D. McClenahan [a], Christine Uhlenhaut [b], Philip R. Krause [a,*]

[a] Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, MD 20892, USA
[b] Highly Pathogenic Viruses ZBS1, Centre for Biological Threats and Special Pathogens, Robert Koch Institute, Nordufer 20, 13353 Berlin, Germany

ARTICLE INFO

*Article history:*
Received 25 April 2013
Received in revised form
28 October 2013
Accepted 8 November 2013

*Keywords:*
Virus detection
Massively parallel sequencing (MPS)
Non-specific PCR

ABSTRACT

Massively parallel sequencing (MPS)-based virus detection has potential regulatory applications. We studied the ability of one of these approaches, based on degenerate oligonucleotide primer (DOP)-polymerase chain reaction (PCR), to detect viral sequences in cell lines known to express viral genes or particles. DOP-PCR was highly sensitive for the detection of small quantities of isolated viral sequences. Detected viral sequences included nodavirus, bracovirus, and endogenous retroviruses in High Five cells, porcine circovirus type 1 and porcine endogenous retrovirus in PK15 cells, human T-cell leukemia virus 1 in MJ cells, human papillomavirus 18 in HeLa cells, human herpesvirus 8 in BCBL-1 cells, and Epstein–Barr Virus in Raji cells. Illumina sequencing (for which primers were most efficiently added using PCR) provided greater sensitivity for virus detection than Roche 454 sequencing. Analyzing nucleic acids extracted both directly from samples and from capsid-enriched preparations provided useful information. Although there are limitations of these methods, these results indicate significant promise for the combination of nonspecific PCR and MPS in identifying contaminants in clinical and biological samples, including cell lines and reagents used to produce vaccines and therapeutic products.


## 1. Introduction

The availability of massively parallel sequencing (MPS)-based virus detection methods has dramatically increased the pace at which new viruses are discovered. Discoveries in the past few years have included Merkel cell polyomavirus [1], Dandenong virus, a novel arenavirus involved with a cluster of fatal transplant-associated diseases [2] and Bas-Congo virus, a novel rhabdovirus associated with hemorrhagic fever [3]. The potential utility of these methods in a regulatory setting was demonstrated when porcine circovirus type 1 (PCV-1) sequences were found in a U.S.-licensed rotavirus vaccine using an MPS-based method [4]. Despite the great power of these methods to identify viral sequences, they are not routinely used in testing of vaccines produced in cell culture, in part due to lack of standardization and uncertainty regarding how to follow up on potential positive results. Although these methods generally involve a combination of virus enrichment, nonspecific PCR amplification, and massively parallel sequencing, different investigators have used varying approaches.

Adventitious agent detection methods are currently employed throughout the virus vaccine production process. For vaccines produced in cells derived from Metazoa (including birds, mammals, and insects), adventitious agent testing may be performed on cell banks, on reagents used in the production process, on drug substance (for viral vaccines, typically corresponding to bulk harvests), and on drug product (corresponding to the vaccine as formulated for administration). While improved adventitious agent testing could potentially be of value at any of these stages, experience has shown that the cell bank is among the more likely stages at which adventitious agents may exist and be undetected. For example, SV40 contamination of early polio vaccines arose because SV40 was present in the primary rhesus monkey kidney cells used to produce vaccines [5,6], and the PCV-1 in rotavirus vaccines was traced to Vero cell banks used to produce the vaccine [7]. Thus, we sought to evaluate approaches that could be suitable for evaluating cell banks for the presence of potential adventitious viruses.

To help address the question of how these next-generation sequencing methods could be used to qualify cell substrates, we used representative DNA and RNA samples and banks prepared from cells known to contain constitutively expressed viruses. We then used these samples to assess components of virus detection methods, including virus enrichment, nonspecific PCR amplification schemes, and MPS methods.

* Corresponding author. Tel.: +1 301 796 1862.
*E-mail address:* philip.krause@fda.hhs.gov (P.R. Krause).

1045-1056/$36.00 © 2013 Published by Elsevier Ltd on behalf of The International Alliance for Biological Standardization.
http://dx.doi.org/10.1016/j.biologicals.2013.11.002

## 2. Materials and methods

### 2.1. Samples

We used cell lines constitutively expressing virus particles or nucleic acids, including human cell lines MJ (ATCC CRL-8294), BCBL-1, Raji (ATCC CCL-86) and HeLa (ATCC CCL-2), porcine kidney PK-15 (ATCC CCL-33), and High Five insect cells (Invitrogen; BT1-TN-5B1-4). Each cell line contained known viruses or viral nucleic acids (Table 1).

The mammalian cell lines were grown in 75 $cm^2$ flasks at 37 °C in 5% $CO_2$ with 5% fetal bovine serum and DMEM or MEM (Invitrogen). Insect cells were maintained in 75 $cm^2$ flasks at 27 °C with serum-free Express Five medium (Invitrogen). Cells and supernatants were harvested, the cells were counted, and aliquoted to contain $1 \times 10^6$ cells/ml.

MS2 phage viral RNA (Roche) was used in 10-fold serial dilutions as templates to examine the sensitivity of DOP-PCR-based nonspecific amplification.

### 2.2. Nucleic acid extraction

Aliquots of the positive control cell lines containing $1 \times 10^6$ cells were pelleted at 8000 rpm for 5 min. Supernatants were removed and the cell pellets were frozen and thawed for 3 cycles on dry ice, and disrupted by QIAshredder column (Qiagen, Valencia CA). For some samples (those denoted "direct extraction" or DE), total cellular DNA and RNA were extracted from the cell pellets with the All Prep Kit (Qiagen) according to the manufacturer's instruction, either manually or with the QIAcube (Qiagen). Following extraction, the RNA was treated according to the manufacturer's instructions with DNase (RQ1 RNase-free DNase I, Promega, Madison, WI) to remove any residual DNA.

For other samples, a capsid enrichment (or "capsid preparation", CP) technique was used to enrich for viral particles and reduce the amount of host DNA in each sample as previously described [8–10]. Briefly, cell pellets were resuspended in 100 μl Tris–HCl pH 7.5 buffer, frozen and thawed three times using dry ice, disrupted by a QIAshredder column and treated with 80 U DNase (Sigma) for 30 min at 37 °C. The reaction was stopped with EDTA (final concentration 41.5 mM), and the digested sample was overlaid onto a 1 M NaCl, 10 mM Tris/HCl pH 7.5 cushion. The viral particles in the sample (which were protected from the DNase due to their protein capsid) were pelleted by ultracentrifugation at 30,000× *g* for 1.5 h at 4 °C. The supernatant was removed and nucleic acids from the pelleted viral particles were extracted with the All Prep Kit according to the manufacturer's instruction, either manually or with the QIAcube.

### 2.3. Degenerate oligonucleotide primer polymerase chain reaction

A degenerate oligonucleotide primer polymerase chain reaction (DOP-PCR) was used to nonspecifically amplify viral nucleic acids present in samples, as previously described [8–10]. Briefly, RNA was reverse transcribed using Superscript II (Invitrogen) with random hexamer primers according to the manufacturer's instructions. The DNA and cDNA from the DE and CP were then nonspecifically amplified by DOP-PCR. To perform DOP-PCR, 10 μl of DNA or cDNA was added to a 40 μl PCR mastermix containing 5 μl 10× PCR buffer (Applied Biosystems), 0.2 μl Brij-35 (3%), 2 μl dNTPs (10 mM; Roche), 0.3 μl AmpliTaq LD polymerase (5 U/μl; Applied Biosystems) and 3 μl 40 μM DOP or IL-DOP primer (Table 2).

The PCR was performed with five nonspecific amplification cycles, followed by 35 traditional amplification cycles. The PCR program consists of an initial denaturation for 5 min at 95 °C, followed by 5 cycles of 1 min at 94 °C, 5 min at 25 °C, slow ramping at 0.1 °C/s to 30 °C, 4 min at 30 °C, slow ramping at 0.1 °C/s to 37 °C, 3 min at 37 °C, slow ramping at 0.1 °C/s to 42 °C, 2 min at 42 °C, slow ramping at 0.1 °C/s to 55 °C, 55 °C for 1 min, 72 °C for 2 min, and 35 cycles of 94 °C for 20 s, 55 °C for 1 min, 72 °C for 1 min with the addition of 1 s per cycle to the extension step, and a final extension step at 72 °C for 10 min.

**Table 1**
Cell lines used as positive controls for optimization of nonspecific amplification and high throughput sequencing techniques.

| Cell line | Species | Virus(es) | Nucleic acid | Virus particles |
|---|---|---|---|---|
| PK15 | Porcine | Porcine circovirus type 1 (PCV-1) | DNA | Yes |
| | | Porcine endogenous retrovirus (PERV) | RNA/DNA | Yes/no |
| High Five | Insect | Nodavirus | RNA | Yes |
| MJ | Human | Human T-cell lymphotrophic virus (HTLV-1) | RNA | Yes |
| BCBL-1 | Human | Human herpesvirus type 8 (HHV-8) | DNA | Yes |
| Raji | Human | Human herpesvirus type 4 (HHV-4) | DNA | No |
| HeLa | Human | Human papillomavirus 18 (HPV-18) | DNA | No |

### 2.4. Sensitivity of DOP-PCR

Serial dilutions of known concentrations of MS2 viral RNA were used as templates for amplification to study the sensitivity of DOP-PCR. Amplified PCR products from the serial dilutions were separated on agarose gels and the banding patterns were visually compared. At least 50 clones of PCR products were also sequenced by Sanger sequencing to verify that the viral nucleic acid was amplified at each dilution.

MPS was used together with DOP-PCR on viral RNA from MS2 phage to further verify the sensitivity of the DOP-PCR. Serial dilutions of MS2 viral RNA were quantified with real-time quantitative PCR [11]. The dilutions were amplified with DOP-PCR and Illumina adaptors with indexes were added by PCR as described in the following section. The prepared libraries were sequenced with the Illumina MiSeq instrument.

### 2.5. Preparation of sample libraries for massively parallel sequencing

Adapters for 454 and Illumina MPS were added to DOP-PCR products by overlapping PCR. The DNA and cDNA were first amplified by DOP-PCR, as described above. The 454 and Illumina adapters with multiplex identifiers (MID) or index sequences were then added by overlapping PCR primers (Table 2). The DOP-PCR products were purified twice by MinElute kit (Qiagen) or Agencourt AmPure XP beads (Beckman Coulter Genomics, Danvers, MA) to remove excess primers and free nucleotides. Five μl of purified PCR product was added to a 50 μl PCR mastermix containing 25 μl 2× HotStart mastermix (Qiagen), and 7 μl MPS overlapping adapter primer (40 μM; Table 2). Multiplexing was performed with either the 454 MIDs 1–8, 10, 11, 13, and 14 (Roche 454), and Illumina indexes 1–12 (Illumina). For some experiments, samples from different cell lines were multiplexed into a single MPS lane.

Hot-start PCR conditions used to add adapters were 95 °C for 15 min followed by 20 cycles of 95 °C for 1 min, 58 °C for 1 min, and 68 °C for 1 min. PCR products were visualized on a 1% agarose gel with ethidium bromide, and with the High Sensitivity DNA chip on the Bioanalyzer (Agilent). The PCR products with MIDs or indexes were pooled, and purified twice (Qiagen MinElute or AmPure XP beads) before sequencing, and again visualized and quantified with the High Sensitivity DNA chip on the Bioanalyzer.

In addition to the DOP-PCR/overlapping primer method, we evaluated addition of adapters using adapter ligation kits and using transposon-based kits.