



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Simultaneous rapid sequencing of multiple RNA virus genomes



John D. Neill^{a,*}, Darrell O. Bayles^b, Julia F. Ridpath^a

^a Ruminant Diseases and Immunology Research Unit, National Animal Disease Center, USDA, ARS, Ames, IA, United States

^b Infectious Bacterial Diseases Research Unit, National Animal Disease Center, USDA, ARS, Ames, IA, United States

A B S T R A C T

Article history:

Received 30 August 2013

Received in revised form 14 February 2014

Accepted 18 February 2014

Available online 28 February 2014

Keywords:

Bovine viral diarrhea virus

Bovine coronavirus

Ion Torrent

Genome assembly

Next generation sequencing

Comparing sequences of archived viruses collected over many years to the present allows the study of viral evolution and contributes to the design of new vaccines. However, the difficulty, time and expense of generating full-length sequences individually from each archived sample have hampered these studies. Next generation sequencing technologies have been utilized for analysis of clinical and environmental samples to identify viral pathogens that may be present. This has led to the discovery of many new, uncharacterized viruses from a number of viral families. Use of these sequencing technologies would be advantageous in examining viral evolution. In this study, a sequencing procedure was used to sequence simultaneously and rapidly multiple archived samples using a single standard protocol. This procedure utilized primers composed of 20 bases of known sequence with 8 random bases at the 3'-end that also served as an identifying barcode that allowed the differentiation each viral library following pooling and sequencing. This conferred sequence independence by random priming both first and second strand cDNA synthesis. Viral stocks were treated with a nuclease cocktail to reduce the presence of host nucleic acids. Viral RNA was extracted, followed by single tube random-primed double-stranded cDNA synthesis. The resultant cDNAs were amplified by primer-specific PCR, pooled, size fractionated and sequenced on the Ion Torrent PGM platform. The individual virus genomes were readily assembled by both de novo and template-assisted assembly methods. This procedure consistently resulted in near full length, if not full-length, genomic sequences and was used to sequence multiple bovine pestivirus and coronavirus isolates simultaneously.

Published by Elsevier B.V. Open access under [CC BY license](https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Next generation sequencing technologies are used to screen environmental and clinical samples for the presence of viral pathogens. This has resulted in discovery of new, previously unknown viruses that are often uncultivable (Victoria et al., 2008; Li et al., 2009; Phan et al., 2012; Reuter et al., 2012). These technologies have allowed screening of samples collected from domestic animals (Shan et al., 2011; Phan et al., 2011; Reuter et al., 2012) and from animals in the wild to determine if they are carriers of potentially zoonotic viral pathogens (Donaldson et al., 2010; Li et al., 2010; Phan et al., 2011; Wu et al., 2012). Additionally, deep sequence analysis was applied to screening of commercially available vaccines for the presence of contaminants and variants (Victoria et al., 2010).

Methods to sequence complete viral genomes have been developed. These include methodologies based on PCR amplification

of viral sequences, both in fragments (Rao et al., 2013) or full-length genome amplification (Christenbury et al., 2010). These have primarily focused on specific, closely related viruses. The Sequence-Independent, Single Primer Amplification (SISPA) procedure was first used to amplify DNA populations through the use of random priming and ligation of adaptors for PCR amplification (Reyes and Kim, 1991). This was modified for amplification of viral sequences from serum to include a step where DNase I was used to first degrade host DNA (Allander et al., 2001). This protocol has since been used in the identification of a novel pestivirus (Kirkland et al., 2007), bocaviruses (Cheng et al., 2010) and identification of viruses associated with cytopathic effect in cell cultures of unknown origin (Abed and Boivin, 2009).

Many virological laboratories, especially those associated with diagnostic laboratories, have freezers that contain large numbers of virus isolates collected over varying time spans. Often these isolates have clinical histories associated with them that detail where and when the isolations were made and descriptions of the disease state of the host. There is a wealth of information in these isolates, but up till now, it has been time consuming and expensive to sequence these viral genomes, often requiring sets of strain-specific

* Corresponding author. Tel.: +1 515 337 7730; fax: +1 515 337 7428.
E-mail address: john.neill@ars.usda.gov (J.D. Neill).

Table 1
Sequences of barcode primers.

Barcode 1	GACCATCTAGCGACCTCCAC ^a
Barcode 2	GGCGGAGCTCTGCAGATATC
Barcode 3	ATCGTCGTCGTAGGCTGTCTC
Barcode 4	GGACAAGCACAGCATAGCCA
Barcode 5	GGCATTGGTCGGCACTTGTT
Barcode 6	GGTAGATAAGCGGTCCGGCTC
Barcode 7	AATCACATAGGCGTCCGCTG
Barcode 8	GGCAGGACCTCTGATACAGG
Barcode 9	GGCACTCGACTCGTAACAGG
Barcode 10	GGTCCAGGCACAATCCAGTC
Barcode 11	GGCAGGTTCAAGCGAGGTTG
Barcode 12	ACGGTGTGTTACCGACGTCC
Barcode 13	GTCTCTGTACGACGGTACAG
Barcode 14	TGCGAACCGTTACGGGTGGA
Barcode 15	GGATAGTCGGCGTGCACCAA
Barcode 16	AACACCTCTGGAACACGCC
Barcode 17	GTGGCAGCTACTCAGACGCT
Barcode 18	TCGCACCTTGCATGGTACGG
Barcode 19	GCCGCTGGTCAGGTAGGAAA
Barcode 20	TGGGACACTCGGGATAGTGG

^a 20mer sequence, random priming barcode oligo has 8 random bases on 3' end.

primers for PCR amplification and sequencing. The protocol outlined in this study provides the means to rapidly sequence archival collections of viruses to study the evolution of specific viruses over time. This information is important in determining the relevance of virus strains used in currently marketed vaccines.

2. Materials and methods

2.1. Viruses and RNA extraction

Viral isolates passaged in cell culture and stored frozen at -80°C were used in this study. The two virus groups sequenced were 21 bovine viral diarrhea virus (BVDV) isolates and 19 bovine coronavirus (BoCV) isolates. After isolation/propagation of the viruses on appropriate host cells, the medium was frozen/thawed twice, cell debris removed by centrifugation, and aliquots frozen at -80°C . Routinely, the libraries were constructed from groups of 10 viruses. The virus stock was thawed at room temperature and 180 μl was mixed with 20 μl of $10\times$ DNase I buffer (100 mM Tris-HCl (pH 7.6), 25 mM MgCl_2 , 5 mM CaCl_2). A cocktail of nucleases was added to degrade host nucleic acids as previously described (Victoria et al., 2008). The cocktail consisted of (per virus sample) 5 units DNase I (Life Technologies, Grand Island, NY, USA), 10 units Turbo-DNase (Life Technologies), 5 units Base-line Zero DNase (Epicenter, Madison, WI, USA), 25 units benzonase (EMD Millipore, Billerica, MA, USA), 2 μg RNase A/5 units RNase T1 (Thermo Scientific, Waltham, MA, USA), and 4 units RNase One (Promega, Madison, WI, USA). The virus stock/nuclease mixture was incubated at 37°C for at least 90 min. The viral RNA was purified using the MinElute Virus spin filter kit (Qiagen, Valencia, CA, USA) according to manufacturer's specifications. The RNA was eluted in 22 μl of ultra-pure, RNase-free H_2O (final yield $\sim 20\ \mu\text{l}$) and immediately placed on ice.

2.2. Random-primed cDNA synthesis and sequence analysis

Twenty microliters of RNA were mixed with 100 pmol of a 28mer primer consisting of 20 nucleotides of a known sequence followed by 8 random nucleotides as previously described (Victoria et al., 2009). Table 1 shows the sequence of the 20 primers used in this study. These primers were developed so that the 20 base known sequence was used for PCR amplification of the library as well as served as a barcode for identifying each viral library following pooling and sequencing. The RNA/primer mix was heated at 75°C for 5 min and immediately placed on ice. First strand

synthesis was done in a reaction mix consisting of 0.5 mM dNTPs, $1\times$ first strand buffer, 2 units RNasin (Promega) and 200 units SuperScript II reverse transcriptase (Life Technologies). The first strand reactions were incubated at 42°C for 1 h, heated at 95°C for 2 min and immediately quenched on ice. The second strand reaction was carried out using Sequenase 2.0 (Affymetrix, Santa Clara, CA, USA) and the nucleotides and primers still present following the first strand reaction (Yozwiak et al., 2012). An equal volume of $2\times$ Sequenase buffer with 4 units of Sequenase per reaction was added on ice. The reaction tubes were placed in a thermocycler that was prechilled to 4°C and were ramped from 4°C to 37°C at $0.1^{\circ}\text{C}/\text{s}$. The reaction was continued at 37°C for 10 min. The tubes were heated to 95°C for 2 min and placed immediately on ice. A second aliquot of Sequenase was added (4 units/reaction) and the ramping and incubation repeated with the exception that the 37°C reaction was done for 30 min. The double stranded cDNA was purified using Minelute PCR spin columns (Qiagen) and eluted with 13 μl of ultra-pure H_2O . The cDNA was amplified by PCR using primers consisting of the first 20 known bases of the 28mer used in the cDNA synthesis (Victoria et al., 2009). The PCR reactions consisted of 10 μl of double-stranded cDNA, $1\times$ Pfx50 buffer, 400 μM dNTPs, 100 pM of appropriate 20mer primer and 5 units of Pfx50 polymerase (Life Technologies). The reactions were cycled at 95°C for 2 min and 35 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 1 min. The number of cycles was empirically determined so as not to over amplify the cDNA, where 35 cycles worked well. The DNA was again purified and eluted in 10 μl ultra-pure H_2O , and 1 μl subjected to Agilent Bioanalyzer analysis (Agilent Technologies, Santa Clara, CA, USA). The percentage of the DNA that was 300–400 bp in length was determined for each reaction as was the total DNA concentration in each sample using PicoGreen (Life Technologies). All ten reactions were pooled so that the 300–400 bp cDNAs from each were present in equimolar amounts. The pooled DNA was placed in an end-repair reaction (Ion Plus Fragment Library kit, Life Technologies) followed by ligation of Ion Torrent sequencing adaptors. The DNA was size fractionated using Agencourt Ampure Beads (Beckman Coulter, Indianapolis, IN, USA) using 1.6 volumes of the Ampure beads and was done according to manufacturer specifications. This step removed DNA of 100 bp and less which eliminated adaptor dimers and unligated adaptors. At this point, two 10-genome libraries, for a total of 20 genomes, were pooled for each sequencing run. The 300–400 bp fraction was isolated from the DNA pool using the Pippin Prep (Sage Science, Beverly, MA, USA). The 300–400 bp DNA was subjected to sequence analysis with the Ion Torrent PGM and standard chemistries (Life Technologies) using the 316 sequencing chip.

2.3. Demultiplexing and virus genome assembly

The raw data files that contained the sequence data from each sequencing run of the PGM were demultiplexed using the 20 base barcode sequence. The raw sequencing reads obtained from the Ion Torrent sequencer were exported as SFF files. Barcode sequences unique to this study were used to populate a custom configuration file that was used in conjunction with the Roche "sfffile" tool to separate the reads into individual barcode-specific SFF files. All the sequences in each barcode-specific SFF file was extracted to a FASTQ file by applying the "sff_extract" utility (http://bioinf.comav.upv.es/sff_extract/index.html) and specifying the "–clip" option to remove the sequencing tags, barcode sequences, and low quality sequences. The "fastx_trimmer" program (part of the Fastx Toolkit; http://hannonlab.cshl.edu/fastx_toolkit/index.html) was used to remove the random primer (8-mer) by trimming an additional eight nucleotides from the 5'-end of each read. As a final post-trimming step, a custom script was used to remove all reads that were less than 40 nucleotides in length. The resulting sequence

files were used to assemble full length genomic sequences using SeqManNGen and were edited with the SeqMan software of the Lasergene 10 package (DNASTar, Inc., Madison, WI, USA) using related viral sequences obtained from GenBank as the assembly reference. Assembled genomic sequences were further edited using Aligner (Codoncode, Inc., Centerville, MA, USA). The sequences from each library were submitted to the NCBI Sequence Read Archive with the BioProject number PRJNA237789, BioSamples SAMN02630193, and SAMN02639534 through SAMN02639572.

3. Results

3.1. Sequencing

The Ion Torrent raw sequence data was trimmed of sequencing adaptors, barcodes, the 8 base random primer sequence, low quality sequences and sequences less than 40 bases. After preassembly processing of the reads, an average of 82% of initial reads remained for downstream analysis. The BVDV and BoCV sequencing runs produced 954,798 and 2,268,798 individual sequences, respectively. The percentage of viral sequences in each barcoded library varied but was as high as 87% (Table 2). The ratio of virus:host sequences was highly dependent on the titer of virus in the stock sample. The number of sequences for each barcoded sample in the library varied, even following pooling at concentrations that should have resulted in equimolar amounts of each sample in the library.

To test the accuracy of the Ion Torrent sequencing, a virus previously sequenced by PCR amplification was included to compare sequences generated by this method and by sequencing PCR amplicons. This virus, a BVDV 1b strain isolated from alpaca (GenBank accession JX297520.1; Table 2, library 3, barcode 10), was assembled from Ion Torrent data and was found to have only 1 base difference from the sequence determined earlier (data not shown). This base was ambiguous in the amplicon sequenced by Sanger chemistry (equal peak height of A and G) but was clearly a G by Ion Torrent sequencing.

The sequences from two libraries were analyzed to determine source of any non-viral sequences. The remaining non-viral sequences were assembled into contigs and were used in BLAST searches of GenBank. This revealed that the non-viral sequences were derived from host nucleic acids, both DNA and RNA. The RNA sequences were primarily ribosomal RNAs but some were from more abundant mRNA transcripts.

3.2. Assembly of genome sequences

Assembly of genomic sequences was done using template-assisted assembly. BVDV and BoCV genomic sequences used as templates for assembly were obtained from GenBank (accession numbers JN380086.1 and EF424617.1, respectively). All of the individual virus library sequences resulted in assembled sequences that spanned at least 94% of the genome from which they were derived (Table 2).

In libraries 1 and 2, BVDV genomic sequences were determined for 20 virus isolates. The percentage of viral sequences for each barcoded library ranged from 34.8 to 87.4%. BVDV assemblies contained no internal gaps but the extreme 5' and 3' termini were not present, with the exception of 3' sequences of four viruses (Table 2). The number of nucleotides missing from the termini varied but ranged from 1 to 98 bases. One virus, library 1, barcode 9, had only 658 viral sequence reads but 94.4% of the genome was assembled. This genomic assembly had 8 internal gaps, ranging from 10 to 219 nucleotides. The other 19 viruses had no assembly gaps and genomic coverages of 99.1–99.9%, depending on variable coverage of the genome termini. The low number of sequences for library 1,

barcode 9 was believed to be a result of a pipetting error during the quantitation or pooling of the 10 genomic libraries.

The results of sequencing of 19 BoCV isolates are listed in Table 2, under libraries 3 and 4. The coronavirus genomic RNA is 2.5 times longer than that of BVDV, roughly 31,000 nucleotides versus 12,300. Thus, it required more sequences overall to assemble the complete genome to the same coverage depth. Additionally, because BoCV typically grows to a lower titer, there were lower ratios of virus:host sequences, ranging from 2.7 to 37.8% of the total sequences. A major difference in genome assembly between BVDV and BoCV was that, with only a few exceptions, both the 5' and 3' terminal sequences were determined for each BoCV isolate. Also, due to the longer length of the BoCV genome, more assembly gaps were observed. However, in all cases, assembled sequences spanned more than 97% of the viral genome. More assembly gaps were observed in individual virus libraries having lower total sequences. The BVDV isolate sequenced in library 3, barcode 10 gave results similar to that observed with BVDV isolates sequenced in libraries 1 and 2.

4. Discussion

The advent of new DNA sequencing technologies has led to the development of improved sequencing protocols for detection and characterization of viruses in environmental and diagnostic samples. Many of these protocols use deep sequencing techniques to detect and identify novel, often uncultivable, viruses. This sequencing protocol was designed to sequence viruses rapidly that were isolated from clinical specimens and archived. Often, a clinical report and history is available for isolates making it possible to begin association of genotype with specific phenotypes and time of isolation. Particularly valuable is that this protocol confers the ability to sequence archival collections of specific viruses rapidly in order to study the evolution of the pathogen over time and determine the genotype of the viruses currently in circulation. This information is useful in determining relevance of vaccine strains based on sequences of currently circulating viruses.

The viruses used for sequence analysis were all isolated and passaged in cell culture. Thus, the abundance of host nucleic acids posed a problem in obtaining sufficient numbers of viral sequences to assemble the viral genomes. It was necessary to pre-digest virus samples with high levels of nucleases before purification of viral RNA. The viral RNA was protected within the intact virions while the host RNA and DNA were degraded. The RNA purification procedure included a protease digestion step aiding in the removal of the nucleases before the viral RNA was released from the virion. The nuclease digestion step aided in reaching as much as 87% of all sequences in a barcoded library being of viral origin. Again, the titer of the virus in the original sample was also important in the number of viral sequences obtained. Viruses that grow routinely to low titers require less than 20 libraries in the sequencing run to ensure sufficient reads in each barcoded library to assemble the genomes.

An important aspect of this procedure was sequence independence, requiring no previous knowledge of sequence or pathogens present in a sample. The use of 20 base barcode sequence with 8 base random nucleotides (Victoria et al., 2008, 2009) served both to prime cDNA synthesis and to differentiate each library following sequencing. The 20 base known sequence was also used for PCR amplification of the individual libraries.

The Ion Torrent platform was chosen because it allowed variation in the depth of sequencing based on the size of the chip that was used in the sequencing run. For the most part, the 316 chip worked well, routinely giving sufficient sequencing depth to assemble the genomes of 20 viruses in each sequencing run. The Ion Torrent is known to have indels in the sequence data, especially associated with homopolymeric runs of bases (Bragg et al., 2013).

Table 2
Sequencing results of four 10 genome libraries.

Barcode	Sequences			Ave. read length (nts)	Missing bases (termini)		Internal gaps (nts)	Percent coverage	Ave. coverage depth
	BVDV	Total	Percent BVDV		5'	3'			
Library 1									
1	7802	8923	87.4	131.1	32	98	0	98.9	92.6
2	17,534	26,986	65.0	133.6	26	0	0	99.8	168.6
3	41,049	61,097	67.2	121.2	23	0	0	99.8	369.4
4	33,228	95,452	34.8	130.9	36	15	0	99.6	367.5
5	50,392	64,068	78.6	131.1	20	2	0	99.8	539.9
6	21,651	32,485	66.6	132.7	29	0	0	99.8	255.0
7	61,750	81,708	75.6	132.1	30	0	0	99.8	676.6
8	50,581	59,226	85.4	125.3	32	34	0	99.5	515.0
9	658	925	71.1	130.6	41	49	10, 19, 42, 43, 79, 83, 100, 219	94.4	8.1
10	51,568	63,659	81.0	127.8	44	0	0	99.6	554.3
Library 2									
11	23,339	38,682	60.3	126.3	43	67	0	99.1	274.4
12	19,152	28,631	66.9	123.3	37	8	0	99.7	216.5
13	18,594	31,087	59.8	122.3	40	2	0	99.7	212.5
14	27,577	70,983	38.8	120.1	3	5	0	99.9	286.0
15	12,798	20,457	62.6	123.4	38	1	0	99.7	136.4
16	39,287	51,997	75.5	119.9	22	0	0	99.8	394.0
17	53,531	68,386	78.3	126.6	31	5	0	99.7	558.7
18	22,571	31,240	72.3	123.7	28	6	0	99.7	244.0
19	28,790	59,170	48.7	127.9	34	7	0	99.7	318.2
20	44,626	58,636	76.1	132.1	37	1	0	99.7	491.5
Barcode	Sequences			Ave. read length (nts)	Missing bases (termini)		Internal gaps (nts)	Percent coverage	Ave. coverage depth
	BoCV	Total	Percent BoCV		5'	3'			
Library 3									
1	4347	39,393	11.0	145.7	95	0	2, 7, 20, 22, 27, 32, 46, 49, 66, 68, 104	98.3	23.2
2	26,659	98,382	27.1	140.4	0	0	0	100	135.4
3	17,589	74,924	23.5	144.2	0	0	16	99.9	91.9
4	36,954	139,754	26.4	141.7	0	0	0	100	185.0
5	15,244	60,400	25.2	145.3	0	0	0	100	82.4
6	50,555	13,803	36.6	136.9	0	0	0	100	246.7
7	10,140	173,472	5.8	142.8	0	0	58, 93	99.5	50.5
8	19,789	52,911	37.4	144.8	0	0	0	100	108.1
9	42,958	113,686	37.8	150.6	0	0	0	100	226.0
10 ^a	5181	191,284	2.7	135.7	53	66	0	99.0	69.6
Library 4									
11	6420	214,458	3.0	138.9	0	0	3, 8, 30, 45, 48, 88, 96, 150, 219, 287	96.9	39.3
12	21,079	170,088	12.4	134.9	0	0	45, 72	99.6	120.3
13	10,146	36,250	28.0	147.7	0	0	16, 61	99.8	59.7
14	29,963	119,097	25.0	140.3	0	0	11, 93	99.7	160.3
15	4055	19,928	20.3	156.4	101	0	10, 12, 19, 39, 46, 60, 66, 68, 80, 183, 229	97.4	26.3
16	14,974	73,805	20.3	149.1	10	0	0	100	82.4
17	20,347	231,561	8.8	156.7	0	0	0	100	117.0
18	36,533	112,996	32.3	139.8	0	0	0	100	190.5
19	52,677	165,881	31.8	159.1	0	0	0	100	301.2
20	9848	42,425	23.2	163.6	0	0	22, 62	99.7	65.2

^a Bovine viral diarrhoea virus, strain Mars.

Use of template assisted assembly of the genomes greatly assisted in producing accurate genome assemblies. Limited numbers of de novo assemblies were done and low numbers of errors were found that were attributed to indels. The accuracy of sequences generated by this method was demonstrated by the comparing sequences to those previously generated by Sanger PCR amplicon sequencing. Only 1 nucleotide difference was detected in sequences generated using the two different methods demonstrating the accuracy of the sequencing using the protocol described in this paper.

The genomic sequences of the BVDV isolates were readily assembled from the sequence data, partially owing to the relatively small size of the genomic RNA. However, the extreme 5' and 3'

terminal nucleotides were difficult to obtain, most likely due to the high degree of secondary structure known to be present in these sequences. This may have made priming of first strand cDNA synthesis near the termini more difficult. With the exception of one virus, all BVDV genomes were assembled with no internal gaps in the sequence. The BVDV genome that was not fully assembled had only 658 total viral sequences, yet this was sufficient to assemble greater than 94% of the genome. The low number of sequences was most likely due to a pipetting error during the pooling of the 10 individual genomic libraries. This illustrates the importance of accurate quantitation and pipetting of the individual libraries in making the DNA pool used in the template preparation for sequencing.

The BoCV assemblies tended to have more assembly gaps, due in part to the longer genomic RNA of these viruses. Conversely, the genome termini of the BoCV were determined more often than the BVDV, the reasons for which are unclear. The assemblies containing the most gaps were those that had lower numbers of total sequences. Still, the sequences spanned greater than 97% of the total genome of all BoCV isolates examined in this study. These gaps were easily filled by either resequencing the library or by PCR amplicon sequencing of selected regions. Routinely, there was sufficient double-stranded DNA to be included in a second pooled library if additional sequence data was necessary.

In addition to sequencing archived viruses, this protocol can readily be applied to diagnostic applications. With modifications of the sample preparation, most clinical samples submitted to diagnostic laboratories can be used for analysis. One difference in a diagnostic application is that complete genome assembly would not be necessary but detection of specific viral sequences would be indicative of the presence of the virus. This would require, at least initially, *de novo* assembly and BLAST analysis to determine if viral sequences were present. Preliminary studies showed that viruses can be successfully detected, and in some cases, near full-length genomic sequences assembled, from serum and fecal samples (data not shown). Additionally, with minor modification of the initial double-stranded cDNA synthesis step, both dsRNA and DNA viruses can be sequenced.

5. Conclusions

A deep sequencing method for rapid determination of genomic sequences of multiple viruses simultaneously is described. The practical application of this protocol was demonstrated by the generation of full-length or near full length genomic sequences of archived pestivirus and coronavirus isolates. With minor modifications, this protocol could be used to sequence multiple dsRNA and DNA viruses. This procedure is sequence independent, requiring no previous knowledge of sequence or pathogens present in a sample. Additionally, this shows promise for the rapid identification of multiple viral pathogens in diagnostic settings and comparison of viruses currently in circulation with those in vaccines.

Acknowledgements

The authors would like to thank Kathryn McMullen, Renae Lesan and Patricia Federico for excellent technical assistance. Additional thanks go to Kerrie Franzen and Mary Lea Killian of the National Veterinary Services Lab, APHIS, USDA for their help in sequencing of the libraries.

References

- Abed, Y., Boivin, G., 2009. [Molecular characterization of viruses from clinical respiratory samples producing unidentified cytopathic effects in cell culture](#). *Viruses* 1, 84–90.
- Allander, T., Emerson, S.U., Engle, R.E., Purcell, R.H., Bukh, J., 2001. [A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species](#). *Proc. Natl. Acad. Sci. U. S. A.* 98, 11609–11614.
- Bragg, L.M., Stone, G., Butler, M.K., Hugenholtz, P., Tyson, G.W., 2013. [Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data](#). *PLoS Comput. Biol.* 9, e1003031.
- Cheng, W.X., Li, J.S., Huang, C.P., Yao, D.P., Liu, N., Cui, S.X., Jin, Y., Duan, Z.J., 2010. [Identification and nearly full-length genome characterization of novel porcine bocaviruses](#). *PLoS ONE* 5, e13583.
- Christenbury, J.G., Aw, P.P., Ong, S.H., Schreiber, M.J., Chow, A., Gubler, D.J., Vasudevan, S.G., Ooi, E.E., Hibberd, M.L., 2010. [A method for full genome sequencing of all four serotypes of the dengue virus](#). *J. Virol. Methods* 169, 202–206.
- Donaldson, E.F., Haskew, A.N., Gates, J.E., Huynh, J., Moore, C.J., Frieman, M.B., 2010. [Metagenomic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat](#). *J. Virol.* 84, 13004–13018.
- Kirkland, P.D., Frost, M.J., Finlaison, D.S., King, K.R., Ridpath, J.F., Gu, X., 2007. [Identification of a novel virus in pigs—Bungowannah virus: a possible new species of pestivirus](#). *Virus Res.* 129, 26–34.
- Li, L., Victoria, J., Kapoor, A., Naeem, A., Shaukat, S., Sharif, S., Alam, M.M., Angez, M., Zaidi, S.Z., Delwart, E., 2009. [Genomic characterization of novel human parechovirus type](#). *Emerg. Infect. Dis.* 15, 288–291.
- Li, L., Victoria, J.G., Wang, C., Jones, M., Fellers, G.M., Kunz, T.H., Delwart, E., 2010. [Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses](#). *J. Virol.* 84, 6955–6965.
- Phan, T.G., Kapusinszky, B., Wang, C., Rose, R.K., Lipton, H.L., Delwart, E.L., 2011. [The fecal viral flora of wild rodents](#). *PLoS Pathog.* 7, e1002218.
- Phan, T.G., Vo, N.P., Bonkoungou, I.J., Kapoor, A., Barro, N., O’Ryan, M., Kapusinszky, B., Wang, C., Delwart, E., 2012. [Acute diarrhea in West-African children: diverse enteric viruses and a novel parvovirus genus](#). *J. Virol.* 86, 11024–11030.
- Rao, P.P., Reddy, Y.N., Ganesh, K., Nair, S.G., Niranjana, V., Hegde, N.R., 2013. [Deep sequencing as a method of typing bluetongue virus isolates](#). *J. Virol. Methods* 193, 314–319.
- Reuter, G., Pankovics, P., Delwart, E., Boros, A., 2012. [Identification of a novel astrovirus in domestic sheep in Hungary](#). *Arch. Virol.* 157, 323–327.
- Reyes, G.R., Kim, J.P., 1991. [Sequence-independent, single-primer amplification \(SISPA\) of complex DNA populations](#). *Mol. Cell. Probes* 5, 473–481.
- Shan, T., Li, L., Simmonds, P., Wang, C., Moeser, A., Delwart, E., 2011. [The fecal virome of pigs on a high-density farm](#). *J. Virol.* 85, 11697–11708.
- Victoria, J.G., Kapoor, A., Dupuis, K., Schnurr, D.P., Delwart, E.L., 2008. [Rapid identification of known and new RNA viruses from animal tissues](#). *PLoS Pathog.* 4, e1000163.
- Victoria, J.G., Kapoor, A., Li, L., Blinkova, O., Slikas, B., Wang, C., Naeem, A., Zaidi, S., Delwart, E., 2009. [Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis](#). *J. Virol.* 83, 4642–4651.
- Victoria, J.G., Wang, C., Jones, M.S., Jaing, C., McLoughlin, K., Gardner, S., Delwart, E.L., 2010. [Viral nucleic acids in live-attenuated vaccines: detection of minority variants and an adventitious virus](#). *J. Virol.* 84, 6033–6040.
- Wu, Z., Ren, X., Yang, L., Hu, Y., Yang, J., He, G., Zhang, J., Dong, J., Sun, L., Du, J., Liu, L., Xue, Y., Wang, J., Yang, F., Zhang, S., Jin, Q., 2012. [Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces](#). *J. Virol.* 86, 10999–11012.
- Yozwiak, N.L., Skewes-Cox, P., Stenglein, M.D., Balmaseda, A., Harris, E., DeRisi, J.L., 2012. [Virus identification in unknown tropical febrile illness cases using deep sequencing](#). *PLoS Negl. Trop. Dis.* 6, e1485.