bioRxiv preprint doi: https://doi.org/10.1101/269936; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1	Simultaneous detection of DNA and RNA virus species
2	involved in bovine respiratory disease by PCR-free rapid
3	tagmentation-based library preparation and MinION
4	nanopore sequencing
5	
6	
7	
8	Matthew S. McCabe ¹ *, Paul Cormican ¹ , Dayle Johnston ¹ , Bernadette Earley ¹
9	
10	
11	¹ Teagasc Grange Animal and Bioscience Research Department, Animal & Grassland
12	Research and Innovation Centre, Teagasc, Dunsany, Co. Meath, Ireland, C15 PW93
13	
14	
15	*Corresponding author
16	E-mail: matthew.mccabe@teagasc.ie
17	
18	
19	

20 Abstract

21	The Oxford Nanopore MinION Mk1B is a portable 90 g device that sequences DNA directly
22	at 450 bases/second generating sequence reads in excess of 400 kb. Recent improvements in
23	error rate and speed of library preparation mean that this device has considerable potential for
24	rapid molecular bovine pathogen diagnostics. We tested the MinION for rapid untargeted
25	detection of viral pathogens associated with bovine respiratory disease (BRD), an
26	economically important disease often involving primary infection of the lung by one or more
27	of a number of DNA and/or RNA viruses. We combined three foetal lung cell cultures which
28	were infected with either Bovine Respiratory Syncytial Virus (BRSV), Bovine Herpes Virus
29	1 (BoHV1) or Bovine Parainfluenza Virus 3 (BPI-3). BoHV1 is a DNA virus and BPI-3 and
30	BRSV are RNA viruses. The cell cultures were treated with DNase and RNase to deplete
31	bovine nucleic acid prior to viral nucleic acid extraction and double-stranded cDNA
32	synthesis. Sequencing libraries were generated by PCR-free tagmentation in under 10
33	minutes and loaded onto a MinION sequencer. Approximately 7,000 sequencing reads were
34	generated and analysed using high-throughput local BLAST against the NCBI nr/nt database.
35	The top BLAST hit for 2,937 of these reads was identified as a virus. Of these, 2,926 (99.6%)
36	were correctly identified either as BoHV1, BRSV or BPI-3.
37	

38 Introduction

39	Bovine respiratory disease (BRD) is the most costly disease of beef cattle in North
40	America and causes significant losses in most other cattle producing regions including
41	Ireland [1-3]. It is thought that the majority of BRD cases involve primary infection of the
42	lower respiratory tract (LRT) with one or more viruses which predisposes the LRT to
43	secondary infection with a single or multiple bacterial species [4,5].
44	As viruses are the most common primary pathogen in BRD, vaccines containing BRD
45	associated modified live viruses (MLVs) are widely used in the cattle industry. The MLVs
46	that are included in the most recent BRD vaccines (e.g. Rispoval [®] 4 (Merck) and Bovi-Shield
47	Gold [®] 5 (Zoetis)) are Bovine Respiratory Syncytial Virus (BRSV), Bovine Herpes Virus 1
48	(BoHV1), Bovine Parainfluenza Virus 3 (BPI3). These MLVs have been selected based on
49	viruses that have been most commonly isolated and cultured or detected by qPCR from nasal
50	swabs, lung washes and lung lesions from BRD cases in recent decades.
51	The primers and probes used in the qPCR assays are designed against small fragments
51	The primers and probes used in the qr CK assays are designed against small fragments
52	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated
52	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated
52 53	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated viral genomes and individual assays have to be performed for each BRD virus. In the
52 53 54	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated viral genomes and individual assays have to be performed for each BRD virus. In the majority of BRD cases, no virus is detected either by culturing or by qPCR. For example, in a
52 53 54 55	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated viral genomes and individual assays have to be performed for each BRD virus. In the majority of BRD cases, no virus is detected either by culturing or by qPCR. For example, in a recent report ³ qPCR assays for BoHV1, BVDV, BPI3, BRSV and Bovine Corona Virus
52 53 54 55 56	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated viral genomes and individual assays have to be performed for each BRD virus. In the majority of BRD cases, no virus is detected either by culturing or by qPCR. For example, in a recent report ³ qPCR assays for BoHV1, BVDV, BPI3, BRSV and Bovine Corona Virus detected virus in only 8.7% (391 out of 4444) of nasal swab samples from live animals
52 53 54 55 56 57	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated viral genomes and individual assays have to be performed for each BRD virus. In the majority of BRD cases, no virus is detected either by culturing or by qPCR. For example, in a recent report ³ qPCR assays for BoHV1, BVDV, BPI3, BRSV and Bovine Corona Virus detected virus in only 8.7% (391 out of 4444) of nasal swab samples from live animals displaying BRD symptoms. The viruses used in vaccines (i.e. BoHV1, BVDV, BPI3, BRSV)
52 53 54 55 56 57 58	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated viral genomes and individual assays have to be performed for each BRD virus. In the majority of BRD cases, no virus is detected either by culturing or by qPCR. For example, in a recent report ³ qPCR assays for BoHV1, BVDV, BPI3, BRSV and Bovine Corona Virus detected virus in only 8.7% (391 out of 4444) of nasal swab samples from live animals displaying BRD symptoms. The viruses used in vaccines (i.e. BoHV1, BVDV, BPI3, BRSV) were detected in only 5.4% (241 of the 4444) samples. In addition, despite the wide use of
52 53 54 55 56 57 58 59	(approximately 200 bp) of genomes of known (i.e. cultured and isolated) BRD-associated viral genomes and individual assays have to be performed for each BRD virus. In the majority of BRD cases, no virus is detected either by culturing or by qPCR. For example, in a recent report ³ qPCR assays for BoHV1, BVDV, BPI3, BRSV and Bovine Corona Virus detected virus in only 8.7% (391 out of 4444) of nasal swab samples from live animals displaying BRD symptoms. The viruses used in vaccines (i.e. BoHV1, BVDV, BPI3, BRSV) were detected in only 5.4% (241 of the 4444) samples. In addition, despite the wide use of MLV BRD vaccines, BRD remains the leading cause of natural death in the cattle industry.

63	Recently viral metagenomics using next generation sequencing (NGS) platforms have
64	been employed to survey viruses associated with BRD. This untargeted approach detects
65	known and unknown viruses in a single universal assay. Ng et al. (2015) conducted viral
66	metagenomic analysis of nasopharyngeal and pharyngeal recess swabs from animals with
67	severe BRD symptoms and detected Bovine Adenovirus 3, Bovine Adeno-associated Virus,
68	Bovine Rhinitis A Virus BSRI4, Bovine Rhinitis B Virus, Bovine Influenza D Virus, Bovine
69	Astrovirus, Picobirnavirus, Bovine Parvovirus 2, and Bovine Herpesvirus 6. Bovine
70	Adenovirus 3, Bovine Rhinitis A Virus BSR14 and Bovine Influenza D Virus BRSI1 were
71	the only viruses that were significantly associated with BRD. None of the viruses that are
72	included in current BRD vaccines were detected [7].
73	So far, BRD viral metagenomic studies have used Illumina sequencing by synthesis
74	(SBS). This requires amplification of cDNA using techniques such as multiple displacement
75	amplification (MDA), and library preparation that can take several hours. The read lengths on
76	the Illumina MiSeq are short with a maximum of just 300 bp paired end on a 600 cycle
77	reagent cartridge. SBS is relatively slow with a 600 cycle MiSeq run taking approximately 56
78	hours and no sequence data is available from a MiSeq until the run has completed. However,
79	Illumina SBS platforms remain unsurpassed in terms of number and quality of sequence
80	reads.
81	The MiSeq, which is one of the smaller Illumina NGS platforms (686 mm \times 523 mm
82	\times 565 mm, 572 grams), is not designed for rapid diagnostics in the field. In contrast, the
83	Oxford Nanopore Technologies MinION Mk1B is a pocket-sized (105 mm \times 23 mm \times 33
84	mm, 87 grams) field-deployable sequencing device that is based on nanopore sequencing.
85	DNA is sequenced directly when it passes through recombinant E.coli CcsG nanopores that
86	are embedded in a membrane and each base causes a characteristic change in the membrane

87 current. This allows extremely rapid direct sequencing of individual DNA molecules. In

88 October 2016 the R9.4 MinION flowcell was released which can run at 450 bases/second per 89 nanopore and generate 10 gigabases of data per MinION flow cell. The library preparation 90 with the 'Rapid Sequencing Kit' takes approximately 10 minutes. Loading the library takes 91 approximately 20 minutes and thousands of long sequence reads (maximum length is 92 typically >65 kb) are available for analysis within minutes of loading the library. Unlike the 93 Illumina SBS platforms, the quality of nanopore reads does not decline with length. However, 94 the error rate of reads from the MinION is still far higher than those of Illumina SBS 95 platforms such as the MiSeq.

96 Several labs have used the MinION to detect viruses. Greninger et al (2015) reported 97 untargeted metagenomic detection of high titre Chikungunya Virus (CHIKV), Ebola Virus 98 (EBOV), and hepatitis C virus (HCV) from four human blood samples by MinION nanopore 99 sequencing [8]. However, to obtain $\geq 1 \mu g$ of metagenomic complementary DNA (cDNA) for 100 the library required for the nanopore sequencing protocol, randomly amplified cDNA was 101 generated using a primer-extension pre-amplification method (Round A/B). Briefly, in Round 102 A, RNA was reverse-transcribed with SuperScript III reverse transcriptase using Sol-103 PrimerA, followed by second strand synthesis with DNA polymerase with Sol-PrimerB and 104 PCR amplification (25 cycles). Libraries were then prepared by end repair, adenylation and 105 adapter ligation. After nanopore sequencing on the MinION they obtained sequences with an 106 individual error rate of 24%. Despite the high error rate this allowed identification of the 107 correct viral strain in all four isolates, and 90% of the genome of CHIKV was recovered with 108 97–99% accuracy. Quick et al. (2016) identified high titre Ebola Virus in samples submitted 109 less than 24 h after collection, with a targeted sequencing approach based on Ebola specific 110 PCR primers that took 15-60 min [9]. For this they used targeted reverse transcriptase PCR 111 (RT–PCR) to isolate sufficient DNA for sequencing using a panel of 38 primer pairs that 112 spanned the Ebola Virus genome. Recently, Kilianski et al. (2016) generated unamplified

113 RNA/cDNA hybrids from nucleic acid extracted from either Venezuelan Equine Encephalitis 114 Virus or Ebola Virus cell culture (both RNA viruses) and sequenced them individually on 115 separate sequencing runs on a MinION. They were able to correctly identify each of the RNA 116 viruses following alignment to the respective viral genomes [10]. 117 In the present study, we tested the potential of untargeted nanopore sequencing on the 118 MinION Mk1B for rapid simultaneous identification of a mixture of DNA and RNA viruses 119 that are associated with BRD. Often more than one virus (with either a DNA or RNA 120 genome) causes infection in the respiratory tract of an animal. We sequenced nucleic acid 121 that we extracted from a mixture of three foetal lung cell cultures which were infected either 122 with BRSV, BPI-3 or BoHV1 [11]. BoHV1 is a member of the family Alphaherpesviridae 123 with a 150 kb linear double-stranded DNA monopartite genome. BRSV and BPI-3 are both 124 members of the family Paramyxoviridae each with a 15 kb negative sense single-stranded 125 RNA genome. BoHV1, BRSV and BPI-3 genomes are all packaged in a protein capsid which 126 is surrounded by an outer lipid membrane envelope. BoHV1 also has a protein tegument 127 between the capsid and the envelope. 128 We report correct simultaneous identification of combined DNA and RNA virus 129 species involved in BRD by PCR-free rapid (10 min) tagmentation-based library preparation 130 and nanopore sequencing on the portable Oxford Nanopore Technologies MinION Mk1B 131 sequencer.

132

133 Methods

¹³⁴ Viral cultures

Foetal lung cell cultures, infected with either BoHV1, BPI-3 or BRSV, were stored at -80°C.
The cultures were sourced from (Agri-Food and Biosciences Institute (AFBI), Veterinary

137	Science Division, Stormont, Belfast, N. Ireland). BPI-3 (TCID50 = $10^{6.5}/100 \ \mu$ L in FCL) was
138	cultured from a diagnostic lung sample. BHV-1 (TCID50 = $10^{6.75}/100 \ \mu$ L in FCL) was
139	cultured from a diagnostic lung sample. BRSV RISP_RS SP.C 11/03/15) (TCID50 =
140	$10^{3.75}/100 \ \mu$ L in FCL) is a vaccine strain of BRSV cultured in the FCL cell line.
141	

142 Nuclease treatment

143 The three frozen viral cultures were crushed to a fine powder with a sterile pestle and mortar

under liquid nitrogen. Crushed frozen powder for each virus culture was weighed (BoHV1

145 (480 mg), BRSV (370 mg), BPI-3 (150 mg)) and combined in a 1.5 mL Eppendorf DNA low

bind tube (Eppendorf, Hamburg, Germany). The volume was adjusted to 1 mL with

147 DNA/RNA/DNase/RNase-free PBS (Sigma Aldrich Ltd., Arklow, Ireland) and 2.5 μL

148 RNaseA (4 mg/mL) (Promega, Madison, WI, USA), 100 μ L of 10× Turbo DNase buffer and

149 $10 \,\mu\text{L}$ of Turbo DNaseTM (Life Technologies Ltd, Paisley, UK) were added. The solution was

inverted slowly six times and incubated for 30 min at 37° C. Then another 10 μ L of Turbo

151 DNase was added, the mixture was inverted slowly again 6 times and incubated for a further

152 30 min.

153

154 Nucleic acid extraction

155 Immediately following nuclease treatment, the remaining nucleic acids were extracted with

the QIAamp Ultra Sens Virus Kit DNA extraction kit (Qiagen, UK) according to

157 manufacturer's instructions except that carrier RNA was substituted with 5.6 µL of a solution

158 of 5 mg/mL linear acrylamide (Thermo Fisher Scientific, MA, USA). The final elution was

performed with $2 \times 30 \ \mu L$ of buffer AVE (total 60 μL) which was supplied in the kit.

160

161 Double-stranded cDNA synthesis

162 Double-stranded cDNA was generated with the Maxima H Minus Double-Stranded cDNA

163 Synthesis Kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the

164 manufacturer's instructions. Briefly, to generate the first cDNA strand (i.e. reverse

transcription), 13 μ L of extracted nucleic acid was added to 1 μ L of random hexamer and

166 mixed by gentle pipetting six times then incubated for 5 min at 65° C. The reaction was placed

167 on ice for 1 minute, then centrifuged in a minifuge for 30 sec and placed back on ice. Then 5

168 μ L of 4× First Strand Reaction Mix and 1 μ L of First Strand Enzyme Mix were added, mixed

by slowly pipetting the entire volume up and down six times, then incubated on an Eppendorf

170 Master Cycler (Eppendorf, Hamburg, Germany) at 25°C (10 min), 50°C (30 min) and 85°C (5

171 min). The tube was removed from the Master Cycler and centrifuged for 10 seconds in a

172 minifuge and placed on ice. To generate the second cDNA strand, the entire first strand

173 cDNA reaction mixture (20 μ L) was combined with molecular grade water (55 μ L) (Sigma-

174 Aldrich, Arklow, Ireland), $5 \times$ second strand reaction mix (20 µL) and second strand enzyme

175 mix (5 μ L). The entire volume (100 μ L) was mixed by slowly pipetting up and down six

times and was then incubated at $16^{\circ}C$ (60 min). The reaction was stopped by adding 6 μ L of

177 0.5 M EDTA (pH 8.0) (Sigma-Aldrich, Arklow, Ireland) and pipetting the entire volume up

and down six times. To remove residual RNA, $10 \,\mu L (100 \,\text{U})$ of RNase1 (supplied in the

179 Maxima H Minus Double-Stranded cDNA Synthesis Kit) was added to the second strand

180 reaction mixture which was mixed by slowly pipetting up and down six times and incubated

at room temperature (22-25°C) for 5 min. The double-stranded cDNA reaction was purified

using a Qiagen MinElute PCR clean up kit (Qiagen, Manchester, UK). This removed

183 enzymes and primers and retained purified double-stranded cDNA and gDNA.

184

185 Double-stranded cDNA/gDNA library preparation

186	The purified double-stranded cDNA/gDNA was tagmented using a Rapid Sequencing Kit
187	SQK-RAD001 (Oxford Nanopore Technologies, Oxford, UK). Briefly, 7.5 μ L of double-
188	stranded cDNA/DNA was added to 7.5 μL of FRM (Oxford Nanopore Technologies, Oxford,
189	UK), then mixed by slowly pipetting up and down six times. The reaction was incubated in a
190	thermal cycler at 30°C for 1 min then at 75°C for 1 min. 1 μ L of RAD (Oxford Nanopore
191	Technologies, Oxford, UK) was added to the 15 μL tagmentation reaction and mixed by slow
192	pipetting, then 0.2 μ L of blunt/TA ligase master mix (New England BioLabs, Ipswich, MA,
193	USA) was added to each tube and mixed again by slow pipetting. The reaction (designated
194	pre-sequencing mix) was then incubated for 5 min at room temperature.
195	
196	Running the library

197 RAD (37.5 μ L), H2O (31.5 μ L) and pre-sequencing mix (6 μ L) were combined in an

198 Eppendorf LoBind tube (Eppendorf, Hamburg, Germany) and mixed by pipetting. The

- 199 resulting 75 µL was loaded onto a Spot-on flowcell (FLO-MIN106) (Oxford Nanopore
- 200 Technologies, Oxford, UK) on a MinION Mk1B (Oxford Nanopore Technologies, Oxford,
- 201 UK) according to the manufacturer's instructions. The flowcell was run for 16 h on
- 202 MinKNOW software (Oxford Nanopore Technologies, Oxford, UK) using the Protocol

203 'NC_48hr_Sequencing run_FLO-MIN105_plus_1D_Basecaller.py'.

204

²⁰⁵ High throughput local BLAST search

- 206 The MinION generated 17,138 FAST5 sequence files which we converted to FASTA files
- using pore tools [12]. This resulted in 7,057 FASTA files that contained sequence reads.
- 208 These sequence reads were then subjected to a local high throughput BLAST search against

- the NCBI nr/nt database using a 24 core processor. The top hit BLAST result for each read
- 210 was used for identification of the 7,057 sequences.
- 211

212 **Results**

213 Viral identification

- Results of the high throughput BLAST search of the 7,057 MinION sequence reads are
- summarised in Table 1 and full details are included in S1 Table. A large number of these
- sequence reads (41.6%) were identified as viruses. The vast majority of the virus-identified
- sequence reads (99.6%) were identified as one of the three expected viruses BoHV1, BPI-3
- and BRSV (Table 1). Only 11.6% of the sequence reads were identified as bovine and 46% of
- the sequence reads were identified as non-bovine/non-viral sequences.

220

222 Table 1. Summary of all top virus hits (species level) following high throughput local

223 BLAST search of MinION sequence reads of rapid sequence library prepared from

224 combined BPI-3, BRSV and BoHV1 calf foetal lung cell cultures against NCBI nr/nt

225 database.

Organism identified	Number	% of total	% of virus reads
	of reads	reads	
BPI-3 (RNA virus, 15 kb genome)	746	10.57	25.40
BRSV(RNA virus, 15 kb genome)	139	1.97	4.73
BoHV1 (DNA virus, 150 kb genome)	2,041	28.93	69.5
BoHV (no type designated)	4	0.06	0.14
BoHV5	4	0.06	0.14
Herpesvirus type1	1	0.01	0.03
Bos taurus BCL2/adenovirus	1	0.01	0.03
Nile crocodilepox virus	1	0.01	0.03
Total number of sequence reads	7,057	-	-
Total number of virus sequence reads	2,937	42.2	-

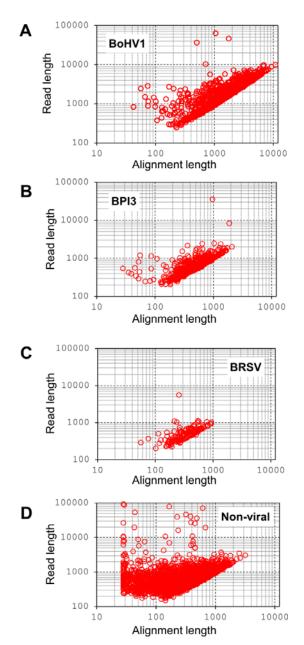
226

227 Read lengths and alignments

Twenty five of the reads were >10 kb and the longest read was 93,542 bp. However, these very long reads only had very short alignments to sequences on the NCBI nr/nt data base (Fig 1). The read length of sequences for which the top BLAST hit was virus, were longer for BoHV1 than for BRSV and BPI-3 (Fig 1). BoHV1 also had the longest alignments out of all viral and non-viral reads (Fig 1). There were three very long viral reads (36,880 bp, 46,607 bp, and 63,234 bp) for which the top hit was either the BoHV1 complete genome or the BoHV1.2 complete genome. However, the alignment bioRxiv preprint doi: https://doi.org/10.1101/269936; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

length of these long reads to these genomes was short (499 bp, 1, 739 bp, 1, 046 bp

respectively) (Fig 1).





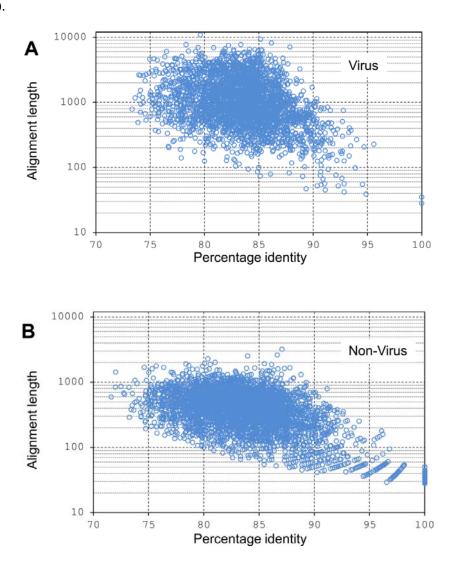
database) against read length of MinION sequence reads. Sequence reads were generated from a rapid

- sequencing library prepared from combined BPI-3, BRSV and BoHV1 calf foetal lung cell cultures. A: Reads
- 241 for which the top BLAST hit was BoHV1. B: Reads for which the top BLAST hit was BPI-3. C: Reads for
- 242 which the top BLAST hit was BRSV. D: Reads for which the top BLAST hit was not a virus.

bioRxiv preprint doi: https://doi.org/10.1101/269936; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

243

The average percentage identity for alignments was 83.5% for viruses and 84.2% for non-viruses (Fig 2). Reads with >90% identity were <2 kb for both viruses and non-viruses (Fig 2).



247

248 Fig 2. Scatterplot of alignment length against percentage identity of MinION sequence reads following

249 high-throughput local BLAST against NCBI nr/nt database. Sequence reads were generated from a rapid

250 sequencing library prepared from combined BPI-3, BRSV and BoHV1 calf foetal lung cell cultures A: Reads for

²⁵¹ which the top BLAST hit was virus B: Reads for which the top BLAST hit was not virus.

Discussion

254	In the present study, combined DNA and RNA viruses were correctly identified following
255	double-stranded cDNA synthesis, library preparation using the Oxford Nanopore
256	Technologies Rapid Sequencing Kit, nanopore sequencing on a MinION Mk1B, and local
257	high throughput BLAST against the entire NCBI nr/nt database. The rapid sequencing library
258	preparation took approximately 10 min and, although the sequencer was run for 16 h, more
259	than 5000 FAST5 files were generated in the first hour of sequencing, which was adequate
260	for identification of the three viruses. This highlights the potential of the MinION for rapid
261	diagnosis of viruses in regional veterinary laboratories and veterinary practices. As the
262	MinION is highly portable, it also has potential for pen-side use by veterinarians and
263	veterinary technicians to characterise viruses in animals that display BRD symptoms and this
264	viral detection method could be applied to DNA extracted from lung washes or
265	nasopharyngeal swabs. However, the numbers of reads generated for the nuclease-treated
266	viral cell cultures was only 7,057 which would be insufficient to detect low titre viruses
267	directly in clinical matrices. With the current depth of sequence available on the MinION, the
268	only way to detect low titre viruses with this device will be through further improvement in
269	depletion of host DNA and unbiased amplification of the library.
270	For nucleic acid extraction and cDNA synthesis we used a spin column method with a
271	microcentrifuge and thermocycler. Magnetic bead-based nucleic acid extraction would
272	eliminate the need for a centrifuge in the field but bead-based extractions tend to be
273	inefficient compared to spin columns. At present, to treat the sample with nucleases, extract
274	the viral nucleic acid then prepare and purify double-stranded cDNA takes approximately 5
275	hours but there is scope to optimise and simplify these steps to significantly reduce this time.
276	We used our own high-throughput BLAST pipeline against the entire NCBI nr/nt data base
277	for 7,057 reads. CPU time for this on a 24 core processor was 130 minutes. We also

278	conducted a local BLAST search against only viral sequences (either against 90,000 virus
279	sequences from NCBI or 7,000 sequences downloaded from Virusite) which took
280	approximately 20 seconds. For virus-only BLAST searches all top hits were viral but many of
281	these hits were not one of the three expected viruses. Reads that did not match the correct
282	viruses were generally less than 100 bp, so removing short reads may work for virus-only
283	BLAST. However, BLAST search against the entire NCBI nr/nt data base, though slower,
284	appeared to be more robust as 99.6% of the viral top hits were correct to viral species level.
285	Due to the low ratio of viral nucleic acid to host nucleic acid, other viral
286	metagenomics protocols commonly use whole genome amplification or PCR. However, in
287	our study amplification was not necessary as 41.6% of the MinION sequence reads were
288	identified as viral by BLAST search against the NCBI nr/nt data base. This may be due to the
289	fact that we used stringent nuclease treatment with Turbo DNase which has $50 \times$ the activity
290	of wild type DNase and that the titre of the virus was high in the cell cultures compared to
291	clinical samples such as swabs. Other protocols also use spin filters to deplete eliminate host
292	cells but it is not clear from the literature whether this is a benefit or an added complication
293	that might bias the viruses that are detected [13]. Therefore we did not include a filtering step.
294	Interestingly the read lengths and alignment lengths were longer for the DNA virus
295	(BoHV1) than the two RNA viruses (BPI-3 and BRSV). This is possibly due to the efficiency
296	of the reverse transcriptase and the DNA polymerase in the double-stranded cDNA synthesis.
297	As such, there may be bias towards DNA viruses in this protocol which will have to be
298	allowed for if accurate quantification of the viruses is required. Direct RNA sequencing
299	(without prior cDNA synthesis) using nanopore technology was recently announced [14].
300	However, for truly universal virus detection by sequencing we require direct DNA and RNA
301	sequencing on the same flowcell which is not currently available.

302	The average error rate for 1D rapid sequencing libraries is currently 15% which
303	explains the low percentage alignments to viral sequences on the NCBI nr/nt database that we
304	observed. Oxford Nanopore Technologies claim this error rate will decrease to 5% for the
305	rapid 1D library preparation and will be 1% for the slower 2D library preps by the middle of
306	2017. Surprisingly, despite the current high 15% error rate and low alignment percentages,
307	our study showed that 99.6% of the MinION sequencing reads, for which the top BLAST hit
308	was a virus, were correctly identified to species level. Several different viral strains were
309	detected for BoHV1. (e.g. BoHV1 strain Cooper, BoHV1 subtype1 and BoHV subtype2).
310	BoHV1 was a field isolate so it is possible that several strains were present. However, with its
311	current error rate it is unlikely that MinION Mk1B 1D rapid sequencing can distinguish
312	between viral strains or subtypes but as error rates are constantly improving this will likely be
313	resolved in the near future.
314	We conclude that double-stranded cDNA synthesis, PCR-free tagmentation, nanopore
315	sequencing on a MinION Mk1B with a 9.4 flowcell, and high-throughput local BLAST
316	search can be used for rapid simultaneous species level identification of mixed RNA and
317	DNA BRD-associated viruses in high titre cell cultures.
318	

319 Acknowledgements

320 We gratefully acknowledge Catherine Duffy and Michael McMenamy (Agri-Food and

321 Biosciences Institute (AFBI), Stormont, Northern Ireland) for provision of viral cultures.

322

References

325	1.	Griffin D. Economic impact associated with respiratory disease in beef cattle. Vet
326		Clin North Am Food Anim Pract. 1997; 13: 367–377.
327		
328	2.	Taylor JD, Fulton RW, Lehenbauer TW, Step DL, Confer AW. The epidemiology of
329		bovine respiratory disease: What is the evidence for predisposing factors? Can Vet J.
330		2010; 51: 1095–1102.
331		
332	3.	O'Neill R, Mooney J, Connaghan E, Furphy C, Graham DA: Patterns of detection of
333		respiratory viruses in nasal swabs from calves in Ireland: a retrospective study. Vet
334		Rec. 2014; 175:351–353.
335		
336	4.	Cusack PM, McMeniman N, Lean IJ. The medicine and epidemiology of bovine
337		respiratory disease in feedlots. Aust Vet J. 2003; 81: 480-487.
338		
339	5.	Murray GM, More SJ, Sammin D, Casey MJ, McElroy MC, O'Neill RG, Byrne WJ,
340		Earley B, Clegg TA, Ball H, Bell CJ, Cassidy JP. Pathogens, patterns of pneumonia,
341		and epidemiologic risk factors associated with respiratory disease in recently weaned
342		cattle in Ireland. J Vet Diagn Invest. 2017; 29: 20-34.
343		
344	6.	Murray GM, O'Neill RG, More SJ, McElroy MC, Earley B, Cassidy JP: Evolving
345		views on bovine respiratory disease: An appraisal of selected key pathogens Part 1.
346		Vet J. 2016; 217: 95–102.

348	7.	Ng TFF, Kondov NO, Deng X, Eenennaam AV, Neibergs HL, Delwarta E. A
349		metagenomics and case-control study to identify viruses associated with bovine
350		respiratory disease. J Virol 2015; 89: 5340-5349.
351		
352	8.	Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, et al. Rapid
353		metagenomic identification of viral pathogens in clinical samples by real-time
354		nanopore sequencing analysis. Genome Med. 2015; 7: 99
355		
356	9.	Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time,
357		portable genome sequencing for Ebola surveillance. Nature 2016; 530: 228–232.
358		
359	10.	Kilianski A, Roth PA, Liem AT, Hill JM, Willis KL, Rossmaier RD et al. Use of
360		unamplified RNA/cDNA-hybrid nanopore sequencing for rapid detection and
361		characterization of RNA virus. Emerg Infect Dis. 2016; 22: 1448–1451.
362		
363	11.	Kirchhoff J, Uhlenbruck S, Goris K, Keil GM, Herrler G: Three viruses of the bovine
364		respiratory disease complex apply different strategies to initiate infection. Vet Res.
365		2014; 45: 20.
366		
367	12.	Loman NJ, Quinlan AR. Poretools: a toolkit for analyzing nanopore sequence data.
368		Bioinformatics. 2014; 30: 3399–3401.
369		
370	13.	Li L, Deng X, Mee ET, Collot-Teixeira S, Anderson R, Schepelmann S, et al.
371		Comparing viral metagenomics methods using a highly multiplexed human viral
372		pathogens reagent. J Virol Methods. 2015; 213: 139-146.

373		
374	14.	Garalde DR, Snell EA, Jachimowicz D, Heron AJ, Bruce M, Lloyd J et al. Highly
375		parallel direct RNA sequencing on an array of nanopores. bioRxiv August 12, 2016
376		
377		
378	Supporting information	
379	S1 Table. Results of local BLAST search of MinION Mk1B sequence reads against the	
380	NCBI	nr/nt database.
381		
382		
383		
384		
385		
386		