RESEARCH REPORT
SWINE HEALTH

Title: Rapid, Pen-side Molecular Diagnostic Tool for Foreign Animal Disease **NPB# 18-040**

Investigator: Cheryl Dvorak

Institution: University of Minnesota, Veterinary and Biomedical Science Department

Date Submitted: July 1, 2020

Industry Summary:

Foreign Animal Diseases (FAD) are black swan events – they come as a surprise, have a major effect, and are often inappropriately rationalized as foreseeable after the fact with the benefit of hindsight. Porcine epidemic diarrhea virus (PEDV) was a wake-up call, but pales in comparison to the consequences of a regulated FAD introduction, such as foot-and-mouth disease virus (FMDV). The challenges addressed in this study were to develop a rapid method for detection of a surprise FAD introduction so that ruinous spread, as occurred with PEDV, and freezing of all animal movement can be avoided. FMDV is particularly challenging since any vesicular disease case must be addressed and FMDV ruled out so that animal transport is permitted. The threat of an undiagnosed FAD introduction, combined with the essential need for rapid rule-outs due to endemic vesicular diseases like Senecavirus A (SVA), requires a robust diagnostic capability. We proposed to develop and benchmark methods and protocols for pen-side sequencing for vesicular disease diagnostics that can provide producers with new options for real-time, robust and rapid molecular detection of foreign and endemic pathogens.

Using SVA as a model for FMDV and PRRSV as a model for endemic pathogens, we evaluated the portable, hand-held MinION sequencing device. The protocols that we developed were compared to the gold standard of detection, diagnostic PCR detection. Our main results showed that MinION sequencing can be used for detection of a pathogen at the species level from a clinical sample without the need for prior knowledge. This can be used for unexpected pathogen detection and co-infection investigations. In addition, our method can improve the detection resolution down to the strain level, which is essential for epidemiological analysis and investigation of other challenging clinical cases, such as a biosecurity breach or vaccine failure. The analytical sensitivity of MinION sequencing as a diagnostic tool was evaluated and compared with real-time PCR. While sequencing at the current state is less sensitive than qPCR diagnostics, it is a great support tool for PCR and provides valuable genomic information.

Pen-side and slaughterhouse nucleic acid sequencing addresses a long-sought goal of moving diagnostic testing into the field for rapid and accurate pathogen identification. It is particularly important for FAD surveillance in suspect cases. We concluded that rapid, direct sequencing can identify pathogens causing clinical disease and in the case of vesicular disease, it was able to identify SVA (even the strain of SVA) as a cause of clinical signs. In addition, this research resulted in a diagnostic method that is immediately adaptable to any pathogen of concern, since it does not require specific primers or specific knowledge of genetic sequence. Thus, it is valuable for a broad range of applications in which timely differential disease diagnosis is needed for case management allowing for improvement in animal care, reduction of costs, and an increase in productivity. Our results have been communicated to the swine community through the SDEC newsletter, the SHMP, the Leman Conference, and AASV.

For more information please contact Dr. Cheryl Dvorak at the Veterinary and Biomedical Sciences Department at the University of Minnesota, email: dvora013@umn.edu.

Key Findings:

- MinION sequencing can be used for robust, unbiased pathogen detection at the species level
- Using additional bioinformatics analysis, the strain of a pathogen can be identified from MinION sequence data
- This study created a disease diagnostic model for FAD outbreaks
- This study accelerated the development of rapid, on-site, real-time disease diagnosis

Keywords: MinION sequencing, rapid diagnosis, bioinformatics, FMDV, SVA, PRRSV

Scientific Abstract:

Foreign animal diseases present a major threat to the global swine industry. FMDV is particularly challenging since any vesicular disease case must be addressed and FMDV ruled out so that animal transport is permitted. In pigs, several other viral vesicular diseases, including swine vesicular disease (SVD), Senecavirus A (SVA), and vesicular exanthema virus infection, cannot be distinguished from FMDV on the basis of clinical findings. The economic implications of the prohibition of pig movement and export due to vesicular disease outbreaks can range from severe for endemic viruses, to catastrophic for FAD. The proposed project thus directly addresses the swine health research priority for rapid, differential diagnostics of vesicular diseases, including FMDV, which is an important FAD. The primary goal is to benchmark rapid sequencing methods for RNA virus detection and evaluate their sensitivities in biological samples that could be collect on the farm or at the slaughterhouse. Rapid, direct sequencing will provide new insights for FMDV detection and differentiation from other pathogens that cause vesicular diseases, such as SVA. Benchmarks will establish key strengths and limitations that must be improved. Our research is directly applicable to all disease pathogens with RNA as the genetic material.

Using SVA as a model, which has been observed as a cause of vesicular disease in different countries throughout the world since 2015, we demonstrated that Oxford Nanopore MinION sequencing could be used as a robust tool for investigation of vesicular diseases. Our results identified the presence of a pathogen from a clinical sample, allowing for identification at the species and strain level. SVA whole genome sequences were generated using both direct RNA sequencing and cDNA-PCR sequencing, with a consensus accuracy of 94% and 99% respectively. The advantages of direct RNA sequencing lie in its simplicity of library preparation and direct RNA strand information which can indicate potential nucleic acid modifications, while cDNA-PCR sequencing excelled at generating highly accurate sequences. This study developed whole genome sequencing methods to facilitate the diagnosis of SVA and provide a reference for investigations of FMD. Next, we used PRRSV, which is the most economically devastating pathogen in the US swine industry, to test the method we established above and to aid in developing more precise detection of endemic diseases, such as examination of co-infection. A nearly full length PRRSV genome was successfully generated from raw sequence reads, achieving an accuracy of 96% after consensus genome generation. Direct RNA sequencing reliably detected the PRRSV strain present with an accuracy of 99.9% using as few as 5 raw sequencing reads and successfully differentiated multiple co-infecting strains present in a sample. In addition, PRRSV strain information was quickly obtained from clinical samples containing 10^4 to 10^6 viral copies or more within 6 hours of sequencing.

Overall, our study not only accelerated the development of robust, rapid, on-site, real-time disease diagnosis but also created a disease diagnostic model for FMD and other FAD outbreaks. Sequencing followed by bioinformatic analysis proves to be a promising approach for investigation of infectious diseases, allowing for more precise prevention and control strategies during outbreaks.

Introduction:

On-farm, pen-side molecular disease diagnosis is increasingly needed and valuable for swine producers, especially when dealing with emerging eventualities such as FAD introduction. Pen-side

detection would provide faster results and facilitate rapid decision-making, an important consideration in an industry in which delays in moving animals due to unknown health status can disrupt flow patterns and schedules, or cause disease outbreaks with great economic losses even in the midst of regional control programs. This study addressed the gap in on-farm molecular diagnosis by evaluating and benchmarking the capacity of the MinION portable sequencer for detection and differentiation of pathogens, with a focus on important foreign animal disease and endemic diseases. The MinION sequencer (Oxford Nanopore Technologies) is a hand-held device that can read sequences >100,000 bases in length. Currently, it only requires about 2h for sample preparation, and generates data in real time to provide real time insights (Juul et al., 2015). MinION sequencing has been used for the detection of Ebola virus, influenza virus, Venezuelan equine encephalitis virus, and other RNA viruses (Hoenen et al., 2016; Kilianski et al., 2016; Wang, Moore, Deng, Eccles, & Hall, 2015).

Foot and Mouth disease (FMD) is one of the most infectious viral diseases known. FMD is important because of its high infectivity and ability to rapidly spread through animal populations and over long distances. Typical clinical signs of FMD are characterized by high temperature, excessive salivation, and formation of vesicles. However, the clinical signs can be confused with other diseases like Senecavirus A (SVA), vesicular stomatitis (VS) and swine vesicular disease (SVD), thus laboratory-based diagnosis for FMD is necessary (Leme et al., 2015). Current laboratory assays for FMDV detection include virus isolation, antigen ELISA and molecular tests (RT-PCR and rRT-PCR) (Jamal & Belsham, 2013). Virus isolation takes several days, is labor-intensive and can only be performed in specialized biocontainment facilities. The limitation of ELISA is that samples with low quantities of virus may yield weak, inconclusive, or negative results (Longjam et al., 2011). Conventional gel-based RT-PCR is not sufficiently sensitive and specific to replace ELISA (House & Meyer, 1993; Longjam et al., 2011). Quantitative RT-PCR assays have greater sensitivity for detection, however, they cannot discriminate between serotypes of FMDV (Rasmussen, Uttenthal, de Stricker, Belak, & Storgaard, 2003). Furthermore, none of these diagnostic methods can detect other vesicular disease pathogens, thus preventing differential discrimination. What is needed is development of a versatile, hand-held FAD detection and diagnosis tool, focusing on FMD and other vesicular diseases, that will provide diagnostic lab sensitivity and specificity without diagnostic lab delays.

MinION sequencing is a robust pathogen detection tool that can be used for metagenomic sequencing to determine the presence and relative abundance of the microbiome in a sample (Greninger et al., 2015; Juul et al., 2015). Compared to a method that can only detect the presence of FMDV, MinION sequencing provides differentiation with other vesicular diseases and extensive additional genetic information about a sample to better inform diagnosis. Traditional sequencing for RNA requires preliminary conversion of RNA to cDNA. Oxford Nanopore sequencing is the only technique that can sequence RNA directly, which eliminates the time and biases introduced during reverse transcription. At present, MinION is mainly used for DNA sequencing, so direct RNA sequencing needs thorough benchmarking to characterize the error profiles and coverage needed for accurate disease diagnosis.

RNA genetic material presents the same technical demands for extraction, processing and sequencing whether it is FMDV, arteriviruses, influenza viruses, coronaviruses, other picornaviruses, rotaviruses, or many foreign animal disease viruses for which rapid pathogen identification and discrimination can be critically important. Knowledge gained from any one virus would be immediately translatable to rapid diagnostic detection and sequencing of an entire class of important swine pathogens. These experiments used PRRSV as model endemic pathogens and SVA as model of FADs due to its molecular similarities to FMDV and genetic diversity that models FMDV diversity. We developed and benchmarked protocols for MinION sequencing from RNA viruses including bioinformatic pipelines to determine the species and strain of virus present as well as determination of analytical sensitivity of the method. These methods can be translated directly into use for detection and analysis of many significant viral diseases of pigs.

Objectives:

Objective 1. Optimize the viral RNA extraction method for MinION sequencing.
Objective 2. Benchmark MinION RNA sequencing for SVA, including error profiles and coverage biases.

Objective 3. Develop a classification model for rapid pathogen detection and differentiation and determine the detection limits (analytical sensitivity) of MinION sequencing for viral detection.

Materials & Methods:

Viral samples

A Senecavirus A (SVA) lab isolate (GenBank: MN164664) and clinical samples of SVA positive vesicular fluid were provided by Dr. Fabio A. Vannucci at the University of Minnesota Veterinary Diagnostic Lab. The SVA lab isolate was propagated in cell culture in NCI-H1299 non-small cell lung carcinoma cell line (ATCC CRL-5803) as previously described (Joshi et al., 2016).

A PRRSV type 2 isolate, VR2332 (GenBank: EF536003.1), was used as the main reference strain. A PRRSV type 1 isolate PRRSV1/USA/Lab6 (SDEU, GenBank: MN175678) and PRRSV type 2 isolate PRRSV2/USA/Lab3 (1-7-4, GenBank: MN175677) were used for experiments examining the detection of multiple viral isolates in a single sample. All viral isolates were propagated on MARC 145 cells as previously described (Robinson, Rahe, Gray, Martins, & Murtaugh, 2018). PRRSV clinical samples and the corresponding ORF5 sequences from Sanger sequencing, were obtained from Dr. Sunil Kumar Mor at the University of Minnesota Veterinary Diagnostic Laboratory (St. Paul, MN).

RNA extraction and viral copy number determination

SVA and PRRSV RNA was extracted from cell culture supernatants, virus-negative pig serum spiked with PRRSV/SVA, and clinical PRRSV-positive serum or SVA-positive vesicular fluid samples using the QIAamp Viral RNA mini kit (Qiagen, Germantown, MD) following manufacturer's instructions without the addition of carrier RNA and with a final elution in 50 ul nuclease-free water. For PRRSV, viral copies were determined using an RT-qPCR assay as described previously using a standard curve to determine the number of viral copies and then calculating the total number of copies sequenced (Robinson et al., 2018). For SVA, the Ct value and viral copies for all samples were determined by RT-qPCR at the University of Minnesota Veterinary Diagnostic Laboratory.

A Qubit 3.0 fluorometer (Life technologies, Carlsbad, CA) and a Nanodrop1000 spectrophotometer (Thermo Scientific, Waltham, MA) were used for quantitative and qualitative assessments.

MinION direct RNA sequencing and PCR-cDNA sequencing

SVA RNA was sequenced using 2 different kits, the direct RNA sequencing kit or the cDNA-PCR sequencing kit (ONT, Oxford, UK). Only direct RNA sequencing was used for PRRSV examination. Direct RNA sequencing (DRS) library preparation was performed according to the direct RNA sequencing online protocol (ONT, Oxford, UK) (Garalde et al., 2018). cDNA-PCR sequencing (PCS) libraries were generated according to the cDNA-PCR sequencing online protocol (ONT, Oxford, UK). After library preparation, the DRS and PCS libraries were loaded onto a R9.4.1 SpotON flow cell and sequenced using a MinION Mk I sequencer (ONT, Oxford, UK) which was connected to a computer and remotely controlled by the MinKNOW software (ONT, Oxford, UK).

For genome sequencing of the cell culture SVA lab isolate, two sequencing replicates were performed for both DRS and PCS, with DRS starting with 60ng SVA RNA plus 300ng RCS (RNA calibration strand) and for PCS 60ng SVA RNA. The RCS is provided with the library kit and was added to the DRS sequencing because more input RNA (~500ng total RNA) is suggested for optimized results. For sequencing runs of clinical samples and negative pig serum spiked with the SVA lab isolate, the same amount of SVA RNA was used for DRS and PCS library preparation, with the addition of 300ng RCS added to each of the DRS samples to increase the amount of total RNA present for optimized sequencing output. Samples were sequenced for approximately 6 hours.

PRRSV cell culture, spike-in and clinical samples were sequenced following the direct RNA sequencing protocol with the addition of exogenous cellular RNA instead of RCS, as needed.

Evaluation of sequencing reads

Basecalling of raw reads was performed using Guppy (Oxford nanopore Technologies Ltd, Oxford, United Kingdom) to generate FASTQ files. Total yield, total reads, read quality, and read length from whole genome sequencing were analyzed using NanoPlot (Van Broeckhoven, Cruts, De Coster, D'Hert, & Schultz, 2018). To obtain raw error rates and error patterns, sequencing reads

were mapped to the reference sequence using minimap2 (Li, 2016), processed with SAMtools (Li et al., 2009) to generate BAM files, and then evaluated by AlignQC (Weirather et al., 2017). Depth of coverage across the consensus genome was analyzed using Qualimap (Okonechnikov, Conesa, & Garcia-Alcalde, 2016). The average coverage and accuracy across the genome were then evaluated and visualized using GraphPad Prism 8 (GraphPad Software, San Diego, CA).

Species and strain detection of SVA

Taxonomic analysis at the species level was performed to identify pathogens existing in spike-in and clinical samples using what's in my pot (WIMP), which is provided by ONT's subsidiary Metrichor (Juul et al., 2015). SVA and PRRSV custom databases were created to analyze SVA/PRRSV sequencing reads by downloading all of the SVA/PRRSV whole genomes from GenBank (132 complete SVA genomes as of March 2019, 949 PRRSV sequences downloaded in Nov 2018). To detect SVA/PRRSV at the strain level, pass reads were analyzed against the SVA/PRRSV custom database using Basic Local Alignment Search Tool (BLAST) (Altschul, Gish, Miller, Myers, & Lipman, 1990) to identify the strain with the best match.

Differentiation of multiple viral isolates in a single sample

Samples containing a mixture of two viral isolates, or VR2332 alone as a control, were sequenced as above. In order to identify the yields needed for accurate strain detection and differentiation, datasets with yields from 30,000 to 30,000,000 bases were generated randomly from total reads using fastq-tools (https://homes.cs.washington.edu/~dcjones/fastq-tools/). PRRSV reads were extracted by mapping all reads to the PRRSV database using minimap2 (Li, 2016). In order to detect PRRSV strains, PRRSV reads were first BLASTn analyzed to identify the top BLAST hit as determined by bit score (BLAST filter of E < 10-50 plus alignment identity > 80% and length > 900 bp). Then, all PRRSV reads were mapped to the this top BLAST hit using minimap2 with "map-ont" preset option (Li, 2016) and mapped reads were extracted using SAMtools (Li et al., 2009). The unmapped reads were also extracted and were analyzed against the PRRSV database a second time to detect any other strain existing in the same sample. The top BLAST hit was recorded and the mapped and unmapped reads to the second top match were again separated. This was repeated until no PRRSV strain was detected in the extracted unmapped reads. The read length and accuracy were based on the results of the analytical sensitivity experiment, where the detection limit was approximately 900 bp and 80% identity. The top BLAST hits were compared to the targeted known strain (1-7-4, SDEU, or VR2332) and the percent identity was recorded. The percentages of reads matching the detected isolates to total PRRSV reads were also recorded.

Consensus generation

For direct RNA sequencing, a consensus genome was generated using the longest PRRSV/SVA read from the sequencing data as a scaffold. The longest SVA or PRRSV read was extracted from the FASTQ file using an awk command, all other raw reads were then mapped to this sequence using minimap2 (Li, 2016) and the map file was processed using Racon (Vaser, Sovic, Nagarajan, & Sikic, 2017). A comparison of this consensus genome to the reference genome was analyzed by pairwise alignment using Geneious software (version 8.0.5) (Kearse et al., 2012). For PCR-cDNA sequencing, de novo assembly was performed using the Canu assembler (Koren et al., 2017).

Using SVA, evaluation and optimization of consensus generation for both sequencing methods (DRS and PCS) was performed in terms of input, total sequence yield, and pre-treatment of raw reads using the cell culture virus in which the whole genome sequence was already known. Groups containing different yields, ranging from 70 kilobases (kb) to 7 megabases (mb), were generated by random selection using fastq-tools (https://homes.cs.washington.edu/~dcjones/fastq-tools/) from the total pass reads dataset. In the same yield group, three subgroups were formed using different raw reads filters; 1) original pass reads without further filters, 2) pass reads with read length > 1314bp to get rid of short reads and RCS in the case of this study, and 3) pass reads that can be mapped to the SVA database. The consensus length and accuracy were the two main parameters evaluated for comparison. Consensus accuracy was determined by a comparison of the consensus genome to the reference genome and was analyzed using the ClustalW pairwise alignment in Geneious 8.0.5 software (https://www.geneious.com) (Kearse et al., 2012).

Evaluation of analytical sensitivity

Analytical sensitivity was determined for both direct RNA sequencing and PCR cDNA sequencing using spike-in and clinical SVA samples containing a range of input virus amounts. For spike-in samples, an SVA viral stock was 10-fold serially diluted from 1X to 10,000X to generate decreasing amounts of virus which were then added into SVA-free pig serum. The Ct value and viral copies for all samples were determined by RT-qPCR at the University of Minnesota Veterinary Diagnostic Lab. For clinical samples, 8 vesicular fluid clinical samples ranging from 10^2 to 10^6 viral copies/ml (Ct values ranging from 24 to 13) were sequenced. For both sample sets, viral RNA was extracted from 1ml of sample with half of the sample used for direct RNA sequencing (DRS) and half for PCR-cDNA sequencing (PCS). For spike-in samples, the SVA strain was determined by blasting raw sequencing reads to the custom SVA database, and then the detected strain was compared to the known reference genome using ClustalW pairwise alignment from the Geneious software to identify the accuracy of the strain level detection. For clinical samples, Sanger sequencing was used to determine a partial 3' sequence (partial 3C, 3D, 3'UTR) which was used as the reference genome sequence. Detection accuracy was determined by comparing the strain identified as the best BLAST match to the reference sequence using ClustalW pairwise alignment in Geneious 8.0.5 software (https://www.geneious.com). A correlation analysis was performed to test if DRS and PCS were quantitative diagnostic methods. The total number of reads varied for each sequencing reaction, thus reads were normalized by calculating the ratio of SVA reads/total reads in order to compare between samples. Linear regression analysis was then performed to determine if there was any correlation using the reads/total reads ratio and the amount of input viral copies using GraphPad prism software (GraphPad Software, San Diego, CA).

Similarly, analytical sensitivity of sequencing for PRRSV was determined using spike-in as well as clinical samples. First, a custom PRRSV sequence database containing 951 PRRSV whole genome sequences was generated by downloading all PRRSV whole genome sequences available in GenBank (949 sequences including our VR2332 strain, download date: Nov 2018) with the addition of sequences from our SDEU and 1-7-4 lab strains. Then, the PRRSV reads were able to be identified and obtained by mapping the raw sequencing reads to this custom PRRSV database using minimap2 (Li, 2016) and extracting the mapped reads using SAMtools (Li et al., 2009). Identification of the viral strain present in the sample was examined using Basic Local Alignment Search Tool (BLAST) with a significance filter of expect value (E) < 10-50 to examine the PRRSV sequence reads. The PRRSV raw reads were compared to the custom PRRSV database using nucleotide BLAST (BLASTn) and the top match, based on bit score, was regarded as the strain detected in the sample. This detected sequence was then aligned to the known reference genome using Geneious software version R8.0.5 (Kearse et al., 2012) and the percent identity was recorded to show the accuracy of detection. For supernatant and spike-in samples, both the VR2332 whole genome and the ORF5 sequence were known and designated as the reference sequence to compare to the MinION generated sequences. For clinical samples, only the ORF5 sequence was known and was used as the reference sequence for comparison. A consensus genome was generated, if possible, for each dataset or sample using the longest PRRSV read as a scaffold followed by analysis of consensus length and accuracy as described above.

Computer codes and sequencing data

The main bioinformatic methods and codes used in this study can be found here: https://github.com/ShaoyuanTan/SVAproject

Results:

Objective 1. Optimize the viral RNA extraction method for MinION sequencing.

Viral RNA was isolated from PRRSV and SVA samples using QIAamp Viral RNA mini kit, which has been shown previously to give good RNA quality and quantity. RNA quantity was determined by Qubit 3.0 fluorometer and RNA quality was examined by a Nanodrop1000 spectrophotometer, targeting at OD260/280 of 2.0 and OD 260/230 2.0-2.2.

Table 1. A comparison of sequencing statistics between viruses and sequencing methods.*

Sequencing statistics	Direct RNA sequencing (SVA)	PCR-cDNA sequencing (SVA)	Direct RNA sequencing (PRRSV)
# available pores	403 ±100	421 ± 18	485 ± 15
Virus yield (Mb)	4.5 ±0.6	66.1 ± 16.2	23.8 ± 4.8
# Virus reads	3559 ± 358	38544 ±0	19255 ± 6070
Mean Virus read length	1267 ±40	1721 ±47	1257 ±146
Virus read error rate (%)	15.14 ± 0.32	11.23 ± 0.23	14.64 ± 0.71

*Data shown as Mean ± SD of 2 independent replicates. SVA sequenced 60ng RNA for 6h and PRRSV sequenced 500 ng RNA for 1h.

Sequence results from SVA samples showed both DRS and PCS can generate long reads over 7000 bp, close to the full-length genome size of 7268 bp. PCS had a much better performance than DRS in terms of higher SVA yield (DRS 4.5Mb, PCS 66.1Mb), longer average read length (DRS 1267bp, PCS 1721bp), and lower raw error rates (DRS 15.14%, PCS 11.23%) (Table 1). Direct RNA sequence results from PRRSV samples gave mean read lengths of approximately 1250 bases, mean error rates of approximately 14.64% and the longest raw read was over 15,000 bp (close to the full-length genome of 15,182 bp) (Table 1).

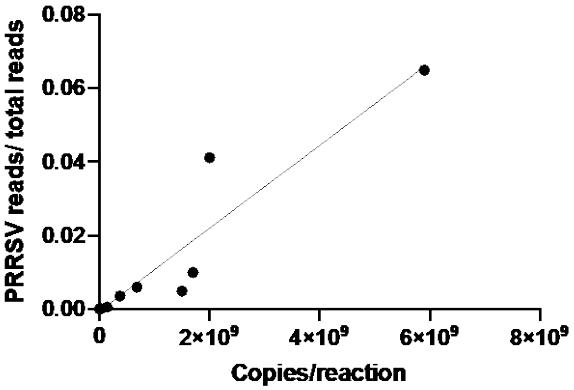

Figure 1. Correlation of PRRSV input copies and raw seqeuncing read output. r^2=0.93, p<0.05.

One problem with the direct RNA sequencing kit was that a large amount of input RNA was needed; much more RNA than would ever be found in a clinical sample. Thus, we examined the addition of an exogenous mRNA or RCS (control RNA provided by Oxford nanopore) sample to the desired RNA sample to increase the RNA quantity to the suggested amount for library preparation and sequencing. An examination of decreasing amounts of input PRRSV genome copies along with added exogenous mRNA was performed and a strong correlation (r^2=0.93, p<0.05) between the number of input viral copies and the yield of viral sequences obtained was observed (Figure 1). Similar findings were observed with SVA samples (r^2 = 0.99 , p < 0.05). Thus, the addition of exogenous mRNA did not deleteriously affect the ability to sequence the virus and knowing the viral copy number in a sample could predict the number of expected sequencing reads and vice versa.

Objective 2. Benchmark MinION RNA sequencing for SVA, including error profiles and coverage biases.

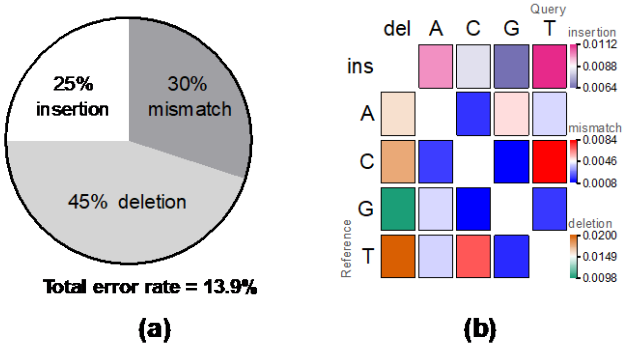

Figure 2. Analysis of direct RNA sequencing errors. To obtain raw error rates and error patterns, raw reads were mapped to the VR2332 reference sequence, followed by evaluation of the mapping. (a) The percent of each error type is shown as well as the total error rate. (b) The error patterns of insertions (first row with darker pink indicating higher errors), deletions (first column with darker orange indicating higher errors), and mismatches (center matrix with darker red indicating higher error). The U bases in the query sequence were adjusted to T automatically by the minimap program in order to map to the reference sequence which was DNA.

A thorough understanding of MinION direct RNA sequencing was performed using PRRSV sequence outputs. First, analysis of PRRSV sequencing results was evaluated for total yield, total reads, read quality, and read length as explained in objective 1 above. The raw error rates and error patterns were examined by mapping to the reference genome sequence. An examination of the error rates showed a total error rate of almost 14% with the highest error coming from deletions (Figure 2a). The most frequently observed error patterns were insertions or deletions of T nucleotides and C and T mismatches (Figure 2b).

Genome coverage was observed to be uneven with high coverage on the 3' end of the genome and low coverage at the 5' end, but consensus accuracy across the genome was similar (93-97%) in spite of the large differences in coverage (Figure 3).

The direct RNA sequencing (DRS) method and analysis was also performed for SVA, showing similar results to that of PRRSV DRS sequncing (Figure 4). In addition, sequencing of the SVA genome was performed using the MinION PCR-cDNA kit (PCS). Results showed the PCS method generated higher quality raw reads, provided a much more even coverage distribution across the genome, and had a higher accuracy than direct RNA sequencing (Figure 4, Table 1).

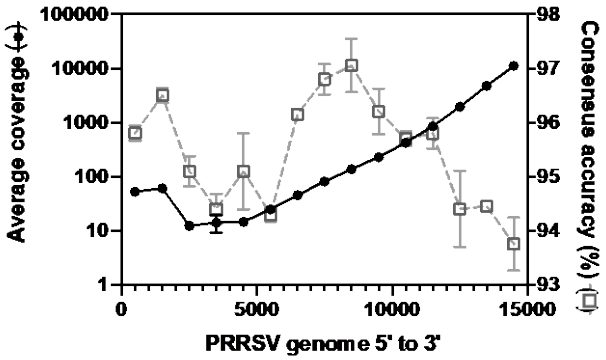

Figure 3. Depth of coverage and consensus accuracy across the PRRSV whole genome. Raw reads were mapped to the longest raw read which served as a scaffold to generate a consensus genome. The coverage sistribution (left y-axis, black closed circles) was evaluated by Qualimap. The consensus accuracy (right y-axis, grey open squares, dashed line) was generated using a pair-wise alignment in geneious software. Both the average coverage and accuracy were evaluated using a window size of 1000bp and visualized using GraphPad prism software.

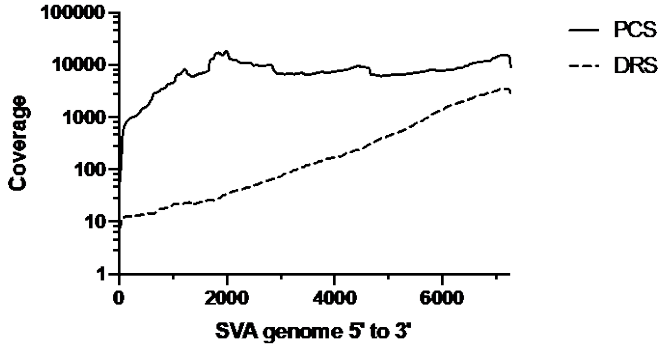

Figure 4. Coverage distribution of direct RNA sequencing and cDNA-PCR sequencing. SVA reads were mapped to a reference genome using minimap2 and analyzed by Qualimap to generate coverage information which was then visualized using GraphPad prism software. The solid black line represents the distribution of cDNA-PCR sequencing reads (PCS) and the dashed line respresents the distribution of direct RNA sequencing reads (DRS).

A consensus genome was generated for DRS sequencing using Racon and PCS sequencing using Canu. Optimization of consensus generation was performed for both DRS and PCS data by examining the sequencing yield and different read filters on the consensus quality. Datasets were generated containing different sequencing yields (0.7, 7, and 70 Mb) by randomly selecting reads from the total pass reads dataset. Within the same yield dataset, three groups were generated based on different filters, with group 1 containing all the pass reads (Phred quality > 7), group 2 consisting of pass reads with a length filter > 1314 bp to remove short reads and all RCS reads (RCS was added to DRS to increase the efficiency of library preparation), and group 3 with pass reads that mapped to the SVA database. Using the "70Mb yield" datasets as an

Table 2. Performance of SVA consensus generation using different raw read filters at different yields.*

		Group 1		Group 2 (length filter)		Group 3 (SVA mapped)		
Method	Yield (Mb)	Consensus length (bp)	Accuracy (%)	Consensus length (bp)	Accuracy (%)	Consensus length (bp)	Accuracy (%)	SVA reads^
DRS	0.7	5098 ± 786	89.3 ± 2.4	4881 ± 1728	89.8 ± 1.8	6057 ± 1143	86.2 ± 3.5	55 ± 2
	7	7155 ± 4	90.8 ± 1.5	5522 ± 1229	91.2 ± 3	7091 ± 45	90.8 ± 1.3	548 ± 20
	70	7163 ± 18	94.3 ± 0.2	7096 ± 30	94.4 ± 0.4	7110 ± 21	94.4 ± 0.5	3559 ± 358
PCS	0.7	6738 ± 424	99 ± 0.1	6316 ± 541	97.4 ± 2.4	6592 ± 489	98 ± 1.0	410 ± 6
	7	7267 ± 132	98.9 ± 0.4	7238 ± 19	99.0 ± 0.1	7079 ± 161	99.0 ± 0.0	4092 ± 64
	70	3053 ± 35	90.5 ± 11.5	6596 ± 1903	91.6 ± 5.4	6761 ± 995	87.9 ± 2.0	38544 ± 10439

*Data shown as Mean ± SD of 2 independent replicates

^# of reads mapped to SVA, only determined for group 3

example, we evaluated the effect of the different filters on read recovery, read length and read quality. Read recovery of the DRS dataset after the length filter (group 2) was 11% of the pass reads (group 2 yield/group 1 yield) and after the SVA mappable filter (group 3) it was 7% of the pass reads (group 3 yield/group 1 yield). The low recovery was due to the large number of short reads present, mainly RCS reads. Read recovery of the PCS dataset after the length filter (group 2) and SVA mappable filter (group 3) was 75% and 73% of the pass reads (group 1), respectively. The average read length and Phred quality score was greater for PCS than for DRS irrespective of the filters. The average read length of the unfiltered DRS reads (group 1) was especially low and was mainly due to the presence of short RCS reads, which account for >90% of reads that are less than 1314bp. Next, examination of each read filter at different sequence yields was performed to determine optimal conditions for generation of a consensus sequence. For the DRS groups, as the starting yield increased, the length and accuracy of the generated consensus sequence increased (Table 2). The highest consensus length and accuracy were observed at the 70Mb yield (~ 5Mb of SVA reads, 7% SVA mappable rate) (Tables 2). At the same yield level, the consensus accuracy and length with different filters were similar, indicating that the sequencing yield is the leading factor for consensus accuracy and length and the raw read filters have minimal influence on results (Table 2). A similar result was observed for PCS sequencing, as within a sequencing yield, the different filters showed similar consensus length and accuracy (Table 2). However, for PCS an increase in yield didn't always result in better consensus generation, as 70 Mb pass reads generated a lower accuracy and a shorter consensus than that of the 7 Mb read group (Table 2). The most accurate consensus for PCS was generated using a total sequencing yield of 7 mb (~5Mb of SVA reads, 73% SVA mappable rate) (Table 2). While both DRS and PCS can generate a nearly full length SVA genome, the consensus from PCS achieved a 99% accuracy, much higher than that of DRS which only reached a 94% accuracy (Table 2). In this study, no obvious differences were observed when comparing different filters using the cell culture samples. However, filters may be useful in some situations not examined here such as in clinical samples from tissues that would contain a large amount of host RNA. In order to make our pipeline applicable to all sample types, we used the SVA mappable reads (group 3 filter set) for the following spike-in and clinical sample analysis.

Objective 3. Develop a classification model for rapid pathogen detection and differentiation and determine the detection limits (analytical sensitivity) of MinION sequencing for viral detection.

In order to detect viruses at the species level in an unbiased and hypothesis-free manner, taxonomic analysis was performed using the what's in my pot (WIMP) workflow. Here, SVA clinical samples were sequenced and analyzed. Results showed that SVA was able to be easily detected using both sequencing methods in both spike-in and clinical samples containing more than 5 x 10^4 total viral copies (Table 3). Using DRS to investigate spike-in samples, SVA reads were detected in samples with as low as 4.7 x 10^2 SVA viral copies, while for PCS, SVA reads were detected in samples with viral copies of 1.2 x 10^4 or greater (Table 3). In clinical samples, the analytical sensitivity for species detection was 9.2 x 10^2 viral copies for DRS and 2.2 x 10^3 viral copies for PCS (Table 3). Other

Table 3. Analytical sensitivity of direct RNA sequencing (DRS) and PCR-cDNA sequencing (PCS).

					Species detection			Strain level identification					Consensus generation					
	Ct value	Viral copies	Total reads		WIMP SVA reads		GenBank reference sequence	DRS		PCS			Minimap SVA reads		DRS consensus		PCS consensus	
			DRS	PCS	DRS	PCS		Best match	Identity (%)	Best match	Identity (%)		DRS	PCS	Length (bp)	Accuracy (%)	Length (bp)	Accuracy (%)
Spike-in samples	10	1.0×10^7	94572	16515	5722	2913	MN164664	MN164664	100	MN164664	100		42936	4658	7189	94.8	7395	99.2
	13	1.2×10^6	40555	5775	236	1182	MN164664	MN164664	100	MN164664	100		1961	1611	7177	91	6873	99
	18	6.5×10^4	185259	5071	17	296	MN164664	MN164664	100	MN164664	100		43	420	2551	90.6	6770	99.4
	20	1.2×10^4	219789	3889	3	3	MN164664	MN164664	100	MN164664	100		10	3	1306	90.3	916[†]	80.8
	25	4.7×10^2	110641	7880	1	0	MN164664	MN164664	100	NA	NA		1	0	456[†]	85.1	NA	NA
Clinical Samples	13	1.1×10^6	45478	78075	299	436	MN990489	KX019804.1	97.1	KX019804.1	97.1		377	630	7157	97	7013	99.6
	16	1.3×10^5	148513	5256	83	7	MN990490	KX019804.1	98.2	KU058182.1	97.8		124	9	3421	96.4	880[†]	89
	18	5.0×10^4	47411	174551	19	24	MN990491	KY618836.1	97.4	KY618836.1	97.4		20	27	6534	97.1	2285[†]	95.3
	20	1.2×10^4	54964	5943	3	0	MN990492	KU051394.1	97.9	NA			4	0	1957	82.4	NA	NA
	21	7.6×10^3	39465	242239	3	4	MN990493	KY618835.1	97.7	MK256736.1	97.4		3	5	1158	94.3	988[†]	87
	22	2.3×10^3	53359	96661	1	4	MN990494	KY618836.1	97.4	KT827250.1	97.7		1	7	511[†]	88.8	7745[†]	85
	23	2.2×10^3	58929	633632	4	1	MN990495, MN997126	KX019804.1	98.2	MH634514.1	98.2		4	1	3206	92.1	1171[†]	94.5
	24	9.2×10^2	41645	135552	1	0	MN990496	MH490944.1	97	NA	NA		1	0	300[†]	80	NA	NA

[†]Longest raw read

observations include the low portion of reads detected as virus (Table 3), which indicates the need to further optimize the sequencing or enrich viral material to increase assay sensitivity.

For epidemiological and precision infection control purposes, it is necessary to know not only the infectious virus present, but the strain of the virus causing disease. Thus, to investigate whether DRS or PCS can identify the strain of SVA that is present in a sample, total reads were BLASTn analyzed against our SVA whole genome database and the sequence with the best match (top BLAST hit based on bit score) was considered as the SVA strain present in the sample (Table 3). The percent identity between the top BLAST hit and the known sequence of the sample was determined to identify the accuracy of strain level detection (Table 3). For the spike-in samples, a laboratory strain with a known whole genome reference sequence (MN164664) was used and this sequence was also present in our SVA whole genome database. The MN164664 sequence was compared to the top BLAST hit to determine the percent identity which indicates the accuracy of strain level detection (Table 3). For each of the clinical samples, a partial genome reference sequence was obtained using Sanger sequencing, which was then used to compare with the top BLAST hit, but since these reference sequences are partial sequences, they are not present in our SVA whole genome database, so we do not expect a 100% identity between the top BLAST hit and our reference sequence (Table 3). Both sequencing methods were 100% accurate when detecting strains for the spike-in samples, in which the reference strain was present in the SVA whole genome database and it was identified as the best match (Table 3). For clinical samples, a comparison of the known partial genome to that of the top BLAST hit showed a sequence identity of 97.0 – 98.2% for both sequencing methods (Table 3). Further examination of sequencing accuracy was determined by creating a consensus genome which was then compared to the known reference sequence to determine the sequencing accuracy. All raw reads that mapped to the SVA database were used to generate a consensus sequence. This consensus sequence (or longest read when no consensus could be generated) was then compared to the known viral reference sequence usually showing higher accuracies from samples with a higher amount of virus and with PCS usually giving higher accuracies than DRS (Table 3).

The analytical sensitivity of MinION direct RNA sequencing was also examined for PRRSV, in addition to SVA. PRRSV was able to be identified at the strain level with 99.4% accuracy using the direct RNA sequencing method if it was at a level greater than approximately 10^4 viral copies (Ct value of 20) in PRRSV negative serum samples that were spiked with known amounts of virus (Table 4). When examining PRRSV positive clinical serum samples from the field, the limit of detection of PRRSV by MinION sequencing increased to about 10^6 viral copies (Table 4).

Table 4. Analytical sensitivity of direct PRRSV RNA sequencing

Sample type	Viral copies/ reaction	# of total reads	Top BLAST match	Identity to ORF5/whole genome %	# of PRRSV reads	Consensus/longest read	
						Length (bp)	Accuracy %
Cell supernatant	5.9×10^9	19,198	KC469618.1	100.0/99.9	1,247	8,282	94.5
	2.0×10^9	23,068	KC469618.1	100.0/99.9	949	7,167	93.8
	1.7×10^9	83,192	KC469618.1	100.0/99.9	831	9,187	93.2
	6.8×10^8	116,698	KC469618.1	100.0/99.9	699	5,975	93.1
	3.7×10^8	118,879	KC469618.1	100.0/99.9	422	6,028	93.7
Spike-in	1.5×10^9	322,778	KC469618.1	100.0/99.9	1,589	15,021	93
	9.4×10^6	300,143	KC469618.1	100.0/99.9	45	3,743	90.5
	3.4×10^4	161,569	CS484777.1	99.0/99.4	3	905*	82.1
Clinical	1.4×10^8	77,468	MF327000.1	96.8/-	42	1,984	90.5
	2.4×10^7	266,120	KX192112.1	97.0/-	16	2,431	88.4
	3.8×10^6	286,680	KT581982.1	94.4/-	6	940*	83.7
	2.3×10^5	201,887	ND	-	0	-	-
	6.5×10^4	240,944	ND	-	0	-	-
	3.2×10^4	307,822	ND	-	0	-	-

ND: not detected

*Longest raw read was used

Unfortunately, the majority of clinical samples are found at lower levels than could be detected by MinION sequencing. However, during a disease outbreak, high levels of virus are present and would easily be able to be examined using MinION sequencing.

An examination of the capability of direct RNA sequencing to differentiate between multiple PRRSV strains present in the same sample, representing clinical scenarios of a suspected PRRSV outbreak in a vaccinated herd, was also performed. Mock samples were created containing VR2332

Table 5. Mapping status of direct RNA sequencing on samples containing multiple viral strains.

Groups	Total yield (kb)	# of total reads	# of PRRSV reads	PRRSV reads / total reads	First match			Second match			Third match		
					Top BLAST match (identity %)	# of matching reads	% of PRRSV reads	Top BLAST match (identity %)	# of matching reads	% of PRRSV reads	Top BLAST match (identity %)	# of matching reads	% of PRRSV reads
Control	30	25	20	80%	KC469618.1 (99.9)	20	100%	ND					
	300	245	210	86%	KC469618.1 (99.9)	208	99%	ND					
	3,000	2,451	2,079	85%	KC469618.1 (99.9)	2,044	98%	ND					
	30,000	24,512	20,819	85%	KC469618.1 (99.9)	20,472	98%	ND					
VR2332 + SDEU	30	38	23	61%	KC469618.1 (99.9)	19	83%	ND					
	300	375	234	62%	KC469618.1 (99.9)	194	83%	CS421743.1 (99.8)	35	15%	ND		
	3,000	3,748	2,281	61%	KC469618.1 (99.9)	1,722	75%	SDEU (100.0)	514	23%	ND		
	30,000	37,478	23,004	61%	KC469618.1 (99.9)	17,610	77%	SDEU (100.0)	4,879	21%	ND		
VR2332 + 1-7-4	30	34	26	76%	JA894280.1 (100.0)	18	69%	ND					
	300	335	272	81%	KC469618.1 (99.9)	224	82%	1-7-4 (100.0)	38	14%	ND		
	3,000	3,351	2,699	81%	KC469618.1 (99.9)	2,287	85%	SDEU (100.0)	14	1%	1-7-4 (100.0)	348	13%
	30,000	33,512	26,917	80%	KC469618.1 (99.9)	22,668	84%	SDEU (100.0)	136	1%	1-7-4 (100.0)	3,633	13%

ND: not detected

virus to represent type 2 modified live virus vaccination (MLV) and either an SDEU viral strain (type 1 field strain) or a 1-7-4 viral isolate (type 2 field strain) to represent concurrent field infection (Table 5). The top hit (which in this case was VR2332) was used to remove viral sequences from that strain and the rest of the raw reads were able to give a good match to the other PRRSV strain that was present in the sample. These observations confirm that even with multiple fairly similar genomes present, they could be identified, if at high enough levels (# PRRSV reads appx 250 in this study), even if they were not known to be present in the sample (Table 5). Thus, confirming that sequencing is an unbiased diagnostic tool with robust advantages. While examining this data we also observed carryover in the MinION flow cell from one experiment to the next when a large amount of input RNA was used even after the suggested procedures of washing and storing flow cells between runs. This suggests a need to keep track of what samples have been used on the flow cell previously in case of unexpected results that may be explained by high concentration samples used in the previous run. In the future, a more effective flow cell wash kit or protocol would be expected to solve this technical problem.

A specific classification model was not necessary for detection of the virus as the raw reads or consensus could be BLAST analyzed against the custom PRRSV or SVA database or the whole GenBank database. In this study, a database containing only PRRSV or SVA whole genome sequences was developed to allow for quicker results than BLAST analysis against the entire GenBank database. If another specific pathogen or group of pathogens was of interest, a related sequence database could quickly and easily be developed accordingly to allow for quicker BLAST results.

Discussion:

SVA-induced vesicular lesions present exactly like the ones caused by foot-and-mouth disease (FMD). FMD is an OIE-listed disease (http://www.oie.int) and a priority foreign animal disease for many countries, including the United States. Currently, the USDA actively investigates farms that show vesicular lesions to rule out the presence of FMDV, even when SVA is suspected (APHIS, 2014). That makes the rapid and accurate detection of infectious agents responsible for

vesicular lesions even more crucial (Segalés, Barcellos, Alfieri, Burrough, & Marthaler, 2017). A diagnostic method is needed for quick identification and confirmation of the presence or absence of FMDV to minimize the economic impact of infection on swine farm production. In this study we used Oxford Nanopore sequencing for SVA detection at both the species and strain level in a research lab setting to develop a robust, on-site sequencing tool for investigation of SVA and to serve as a reference for investigation of other emerging RNA viral diseases. Our general workflow involved several steps. A robust and unbiased species-level detection of pathogens via WIMP where we can get a general picture of all pathogens involved in a sample. Combined with clinical symptoms and professional knowledge, suspected pathogen(s) can be targeted. Strain-level detection of the targeted pathogen(s) was then performed to provide more information about the pathogen for further investigations. We also optimized whole genome generation of SVA to provide support to expand the SVA database for further research of pathogenic mechanisms or disease epidemiology. We concluded that sequencing can be used for investigation of emerging diseases by identifying unexpected pathogens and can be incorporated as a supportive tool for PCR and to aid in primer design of new strains or pathogens.

This study used SVA as a model for emerging diseases and PRRSV as a model for endemic disease. PRRSV is one of the most challenging endemic pathogens to control in the swine industry. It has been reported that both Sanger and Illumina sequencing can accurately detect PRRSV strains present in a sample, but both require transcription of RNA into cDNA followed by PCR amplification prior to sequencing (Han, Xu, Wang, & He, 2019; Zhang et al., 2017). Differing from this, MinION technology offers the option to directly sequence RNA strands for detection of PRRSV strains. Most importantly, the MinION sequencer is cost-effective and easily accessible, without the investment of expensive sequencing and bioinformatics infrastructure. Despite the low raw read accuracy of direct RNA sequencing (~86%), which is the main concern with this technology, PRRSV strains were able to be identified. This accurate strain-level detection, even though the sequence accuracy is low, allows for guidance on determining effective control methods due to the precise detection of the circulating strains present on a farm.

In addition, we provided thorough comparisons for 2 MinION sequencing methods, PCS and DRS, aiming to provide guidance on selection of a sequencing method when in different clinical situations and for different purposes. PCS takes more time but can generate a more accurate consensus, the advantage of which was especially obvious with higher viral copy number samples. DRS is less accurate, just as sensitive, but has unique and promising features such as detection of nucleic acid modifications as confirmed by other studies (Ovcharenko & Rentmeister, 2018; Viehweger et al., 2018). Despite the differences between read accuracy, both methods can detect virus at the strain level correctly using raw reads without the need for any further bioinformatics analysis. This allows anyone to run the assay easily since one of the limitations of sequencing as a diagnostic tool is its high requirement for data management and analysis.

Analytical sensitivity of sequencing was analyzed for both PCS and DRS. When it comes to analytical sensitivity of MinION sequencing diagnostics, it is important to define sequencing time, since longer time means more reads thus higher sensitivity. In our study, we sequenced samples for 6 hours in order to set a shorter turnaround time frame for generation of same day results. During disease outbreaks, time from sample collection to a clinical report should occur as quickly as possible and same day results would be ideal. The analytical sensitivity of DRS and PCS was shown to be similar, approximately 10^2 to 10^4 viral copies. So far, the analytical sensitivity of MinION sequencing is lower than that of diagnostic PCR assays, but PCR is limited by availability of assays for detection of known pathogens. However, it has been indicated that with further development, short-read second-generation sequencing will have a sensitivity to detect viruses that is comparable to PCR assays (Huang et al., 2019). We believe as MinION sequencing continues to develop, its ability for pathogen detection will increase. Even at the current state, MinION sequencing can be used as a support for PCR detection to provide strain information for more effective disease control. At a high enough viral presence, sequencing can be used to generate whole genome sequence and expand the genome database. Of note, we observed that direct RNA sequencing is a quantitative assay while cDNA-PCR sequencing (amplicon sequencing) is not. Similar conclusions were drawn from other research groups. In an influenza virus study, a strong relationship between the viral titer and the proportion of influenza reads using Oxford Nanopore direct RNA sequencing technology was

observed (Lewandowski et al., 2019). However, in a hepatitis B virus (HBV) study using Oxford Nanopore amplicon sequencing, authors concluded that amplicon sequencing was not quantitative and considerable variability in total yields and the proportion of mapped HBV reads between sequencing runs was observed (McNaughton et al., 2019). The reason could be that cDNA-PCR sequencing includes more steps during library preparation, including amplification and PCR product selection which can introduce bias, while the process of direct RNA library preparation is simple and straightforward.

The methods established in this study provide a framework for investigation of other emerging or endemic viral diseases and can be applied to aid real-world problems directly. There are a few limitations to this sequencing method. Our method of species and strain detection largely depend on the genome database, GenBank. While investigating emerging viral diseases that are caused by known viruses expanding to new hosts or geographical regions, GenBank will have the viral genome information. Some previously unknown, newly discovered pathogens won't be detected directly from the databases due to lack of sequence information. However, the information is present, and the emerging pathogen can still be determined by carefully examining the unclassified sequencing reads and determining if they are at all similar to previously identified pathogens.

The current evaluation of this sequencing technology indicates that it can be used successfully along with qPCR for diagnosis of a pathogen, whole genome generation, and strain-level pathogen detection and differentiation. The future goal is to realize on-site infectious diseases investigation using the Oxford Nanopore MinION portable sequencer to allow for quicker diagnosis and facilitation of more rapid decision-making, an important consideration in an industry in which delays in moving animals due to unknown health status can disrupt flow patterns and schedules, or cause disease outbreaks with great economic losses. Knowledge gained from PRRSV and SVA in this study can be immediately translatable to aid in rapid diagnostic detection and strain-specific identification of an entire class of important swine pathogens. This technology can provide a complete readout of RNA viruses and RNA from other pathogens present in a sample without the need for pre-existing knowledge of what might be present (Theuns et al., 2018). In fact, MinION sequencing technology might end up being a useful and affordable diagnostic tool for foreign animal diseases in general.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol, 215*(3), 403-410. doi:10.1016/s0022-2836(05)80360-2

APHIS, U. (2014). Foot-and-mouth disease response plan. *The Red Book*.

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., . . . Turner, D. J. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods, 15*(3), 201-206. doi:10.1038/nmeth.4577

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., . . . Chiu, C. Y. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med, 7*, 99. doi:10.1186/s13073-015-0220-9

Han, G., Xu, H., Wang, K., & He, F. (2019). Emergence of Two different recombinant PRRSV strains with low neutralizing antibody susceptibility in China. *Sci Rep, 9*(1), 2490. doi:10.1038/s41598-019-39059-8

Hoenen, T., Groseth, A., Rosenke, K., Fischer, R. J., Hoenen, A., Judson, S. D., . . . Feldmann, H. (2016). Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool. *Emerg Infect Dis, 22*(2), 331-334. doi:10.3201/eid2202.151796

House, C., & Meyer, R. F. (1993). The detection of foot-and-mouth disease virus in oesophageal-pharyngeal samples by a polymerase chain reaction technique. *J Virol Methods, 43*(1), 1-6.

Huang, B., Jennsion, A., Whiley, D., McMahon, J., Hewitson, G., Graham, R., . . . Warrilow, D. (2019). Illumina sequencing of clinical samples for virus detection in a public health laboratory. *Sci Rep, 9*(1), 5409.

Jamal, S. M., & Belsham, G. J. (2013). Foot-and-mouth disease: past, present and future. *Vet Res, 44*, 116. doi:10.1186/1297-9716-44-116

Joshi, L. R., Mohr, K. A., Clement, T., Hain, K. S., Myers, B., Yaros, J., . . . Schaefer, R. (2016). Detection of the emerging picornavirus Senecavirus A in pigs, mice, and houseflies. *Journal of Clinical Microbiology, 54*(6), 1536-1545. doi:10.1128/JCM.03390-15

Juul, S., Izquierdo, F., Hurst, A., Dai, X., Wright, A., Kulesha, E., . . . Turner, D. J. (2015). What's in my pot? Real-time species identification on the MinION. *bioRxiv*. doi:10.1101/030742

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics, 28*(12), 1647-1649. doi:10.1093/bioinformatics/bts199

Kilianski, A., Roth, P. A., Liem, A. T., Hill, J. M., Willis, K. L., Rossmaier, R. D., . . . Rosenzweig, C. N. (2016). Use of Unamplified RNA/cDNA-Hybrid Nanopore Sequencing for Rapid Detection and Characterization of RNA Viruses. *Emerg Infect Dis, 22*(8), 1448-1451. doi:10.3201/eid2208.160270

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res, 27*(5), 722-736. doi:10.1101/gr.215087.116

Leme, R. A., Zotti, E., Alcantara, B. K., Oliveira, M. V., Freitas, L. A., Alfieri, A. F., & Alfieri, A. A. (2015). Senecavirus A: An Emerging Vesicular Infection in Brazilian Pig Herds. *Transbound Emerg Dis, 62*(6), 603-611. doi:10.1111/tbed.12430

Lewandowski, K., Xu, Y., Pullan, S. T., Lumley, S. F., Foster, D., Sanderson, N., . . . Matthews, P. C. (2019). Metagenomic Nanopore sequencing of influenza virus direct from clinical respiratory samples. *bioRxiv*, 676155. doi:10.1101/676155

Li, H. (2016). Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics, 32*(14), 2103-2110. doi:10.1093/bioinformatics/btw152

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Longjam, N., Deb, R., Sarmah, A. K., Tayo, T., Awachat, V. B., & Saxena, V. K. (2011). A Brief Review on Diagnosis of Foot-and-Mouth Disease of Livestock: Conventional to Molecular Tools. *Vet Med Int, 2011*, 905768. doi:10.4061/2011/905768

McNaughton, A. L., Roberts, H. E., Bonsall, D., de Cesare, M., Mokaya, J., Lumley, S. F., . . . Matthews, P. C. (2019). Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV). *Sci Rep, 9*(1), 7081. doi:10.1038/s41598-019-43524-9

Okonechnikov, K., Conesa, A., & Garcia-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics, 32*(2), 292-294. doi:10.1093/bioinformatics/btv566

Ovcharenko, A., & Rentmeister, A. (2018). Emerging approaches for detection of methylation sites in RNA. *Open Biol, 8*(9). doi:10.1098/rsob.180121

Rasmussen, T. B., Uttenthal, A., de Stricker, K., Belak, S., & Storgaard, T. (2003). Development of a novel quantitative real-time RT-PCR assay for the simultaneous detection of all serotypes of foot-and-mouth disease virus. *Arch Virol, 148*(10), 2005-2021. doi:10.1007/s00705-003-0145-2

Robinson, S. R., Rahe, M. C., Gray, D. K., Martins, K. V., & Murtaugh, M. P. (2018). Porcine reproductive and respiratory syndrome virus neutralizing antibodies provide in vivo cross-protection to PRRSV1 and PRRSV2 viral challenge. *Virus Res, 248*, 13-23. doi:10.1016/j.virusres.2018.01.015

Segalés, J., Barcellos, D., Alfieri, A., Burrough, E., & Marthaler, D. (2017). Senecavirus A:An Emerging Pathogen Causing Vesicular Disease and Mortality in Pigs? *Veterinary Pathology, 54*(1), 11-21. doi:10.1177/0300985816653990

Theuns, S., Vanmechelen, B., Bernaert, Q., Deboutte, W., Vandenhole, M., Beller, L., . . . Nauwynck, H. J. (2018). Nanopore sequencing as a revolutionary diagnostic tool for porcine viral enteric disease complexes identifies porcine kobuvirus as an important enteric virus. *Sci Rep, 8*(1), 9830. doi:10.1038/s41598-018-28180-9

Van Broeckhoven, C., Cruts, M., De Coster, W., D'Hert, S., & Schultz, D. T. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics, 34*(15), 2666-2669. doi:10.1093/bioinformatics/bty149

Vaser, R., Sovic, I., Nagarajan, N., & Sikic, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res, 27*(5), 737-746. doi:10.1101/gr.214270.116

Viehweger, A., Krautwurst, S., Lamkiewicz, K., Madhugiri, R., Ziebuhr, J., Hölzer, M., & Marz, M. (2018). Nanopore direct RNA sequencing reveals modification in full-length coronavirus genomes. *bioRxiv*, 483693. doi:10.1101/483693

Wang, J., Moore, N. E., Deng, Y. M., Eccles, D. A., & Hall, R. J. (2015). MinION nanopore sequencing of an influenza genome. *Front Microbiol, 6*, 766. doi:10.3389/fmicb.2015.00766

Weirather, J. L., de Cesare, M., Wang, Y., Piazza, P., Sebastiano, V., Wang, X. J., . . . Au, K. F. (2017). Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. *F1000Res, 6*, 100. doi:10.12688/f1000research.10571.2

Zhang, J., Zheng, Y., Xia, X. Q., Chen, Q., Bade, S. A., Yoon, K. J., . . . Li, G. (2017). High-throughput whole genome sequencing of Porcine reproductive and respiratory syndrome virus from cell culture materials and clinical specimens using next-generation sequencing technology. *J Vet Diagn Invest, 29*(1), 41-50. doi:10.1177/1040638716673404