

MinION nanopore sequencing of an influenza genome

Jing Wang¹, Nicole E. Moore¹, Yi-Mo Deng², David A. Eccles¹ and Richard J. Hall^{1*}

¹ Institute of Environmental Science and Research, National Centre for Biosecurity and Infectious Disease, Upper Hutt, New Zealand, ² WHO Collaborating Centre for Reference and Research on Influenza, Victorian Infectious Diseases Reference Laboratory, Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia

OPEN ACCESS

Edited by:

William M. McShan,
University of Oklahoma Health
Sciences Center, USA

Reviewed by:

Peirong Jiao,
South China Agricultural University,
China

Allison Gillaspay,
University of Oklahoma Health
Sciences Center, USA

*Correspondence:

Richard J. Hall,
Institute of Environmental Science
and Research, at the National Centre
for Biosecurity and Infectious Disease,
66 Ward Street, Upper Hutt 5018,
New Zealand
richard.hall@esr.cri.nz

Specialty section:

This article was submitted to
Virology,
a section of the journal
Frontiers in Microbiology

Received: 09 May 2015

Accepted: 14 July 2015

Published: 18 August 2015

Citation:

Wang J, Moore NE, Deng Y-M, Eccles
DA and Hall RJ (2015) MinION
nanopore sequencing of an influenza
genome.
Front. Microbiol. 6:766.
doi: 10.3389/fmicb.2015.00766

Influenza epidemics and pandemics have significant impacts on economies, morbidity and mortality worldwide. The ability to rapidly and accurately sequence influenza viruses is instrumental in the prevention and mitigation of influenza. All eight influenza genes from an influenza A virus were amplified by PCR simultaneously and then subjected to sequencing on a MinION nanopore sequencer. A complete influenza virus genome was obtained that shared greater than 99% identity with sequence data obtained from Illumina MiSeq and traditional Sanger-sequencing. The laboratory infrastructure and computing resources used to perform this experiment on the MinION nanopore sequencer would be available in most molecular laboratories around the world. Using this system, the concept of portability, and thus sequencing influenza viruses in the clinic or field is now tenable.

Keywords: influenza, virus, nanopore sequencing, MinION, whole genome sequencing

Introduction

Next-generation sequencing technologies (NGS) are now capable of providing a whole genome sequence for most organisms (Vernikos et al., 2015). As such technologies have become more trusted, they are being applied with increasing frequency in clinical microbiology laboratories to detect and characterize pathogens; replacing the incumbent Sanger-sequencing methods (Bertelli and Greub, 2013; Quinones-Mateu et al., 2014). Sanger-sequencing has certainly been a reliable and robust method which has served molecular biology well for over three decades (McGinn and Gut, 2013). However, it is labor-intensive, slow, and not easily adapted for processing large genomes or large numbers of samples.

The characterisation of influenza viruses has certainly benefitted from whole genome sequencing (WGS) using Sanger-sequencing, providing a tool for surveillance of the highly dynamic genomes of influenza viruses (Ghedini et al., 2005). Such data can assist in vaccine development, detection of anti-viral drug resistance, or the identification of new reassortant influenza viruses. These advances may help to mitigate the morbidity and mortality of influenza pandemics or seasonal epidemics. Latterly, NGS platforms such as the Illumina MiSeq and Life Technologies Ion Torrent have received much acclaim about their ability to provide WGS data for influenza, potentially within 2–3 days of the receipt of a sample (Van den Hoেকে et al., 2015). Such methods often use amplicon sequencing, where each of the eight gene segments are amplified individually or using generic primers (Van den Hoেকে et al., 2015). Primers targeting the highly conserved 5' and 3' termini of each gene segment allow for simultaneous amplification of all segments within a single PCR reaction (Zhou et al., 2009). The use of a specific PCR avoids the issue of concomitant sequencing of host RNA (or DNA) that can occur in a metagenomic approach (Nakamura et al., 2009; Hall et al., 2014; Li et al., 2015). It then becomes possible to obtain WGS for influenza viruses that are present at low concentrations in a sample, as is commonly observed for clinical material.

However, there are also difficulties that attend the use of NGS for sequencing influenza. Such technologies require specialized instrumentation and reagents that are likely to be beyond the remit of diagnostic or regional reference laboratories. Furthermore, although the cost of NGS is decreasing, it is still prohibitive unless large numbers of samples are to be processed. There is clearly a place for a low-cost and portable NGS instrument that can be deployed more widely.

The MinION nanopore sequencing device has the potential to be the world's first mobile sequencer (Quick et al., 2014). It is a small device (10 cm × 2 cm × 3.3 cm; approximately 90 g) and is powered by a computer USB port. Analysis of sequencing data is done in real time by an internet-connected laptop. Another unique feature of the device is the ability to generate very long reads of up to 60 kilobases (Madoui et al., 2015). It has been successfully applied for sequencing whole genomes of *E. coli*, poxviruses and lambda bacteriophage (Quick et al., 2014; Kilianski et al., 2015). It has also been applied to identify antibiotic resistance genes in repetitive regions of the genome of *Salmonella typhi* (Ashton et al., 2015). The MinION may therefore provide new opportunities in infectious disease diagnostics, for example, in the rapid sequencing of viral genomes in the clinic in response to the early phase of an influenza pandemic; or determination of ebola virus genotypes during an outbreak in a remote location. We show that whole genome sequence of a human (and animal) viral pathogen, the influenza virus, can be readily obtained using the MinION nanopore sequencer. These data are compared with an established Sanger-sequencing method and another NGS device, the Illumina MiSeq.

Materials and Methods

Nucleic Acid Extraction

The influenza A/New Zealand/316/2014 (H3N2) isolate was obtained from the culture collection at New Zealand's World Health Organization National Influenza Centre, based at the Institute of Environmental Science and Research (ESR). RNA was extracted from 200 µL of MDCK-SIAT viral culture supernatant using the iPrep Purelink Virus kit (Life Technologies, Carlsbad, CA, USA), with elution into 50 µL of RT-PCR molecular grade water (Ambion, Austin, TX, USA).

Sanger-sequencing

The eight influenza gene segments were first amplified using the RT-PCR protocol published by Deng et al. (2015). Briefly, cDNA was made using uni-12 primer [5'-AGCAAAGCAGG] with ThermoScript RT-PCR system for first-strand cDNA synthesis kit (Invitrogen) as per the manufacturer's instructions. Then 2 µL cDNA was added in 17 separate PCR reactions with Platinum Taq DNA polymerase high fidelity kit (Invitrogen) using gene specific primers that tagged with M13 universal sequences to the 5' end. The PCR program used were as follows: 2 min at 94°C, then 35 cycles of 30 s at 94°C, 30 s at 55°C, and 1 min at 68°C, with a final extension at 68°C for 2 min. PCR amplicons were visualized by E-gel (Invitrogen), followed by ExoSAP IT (GE Healthcare) purification and used for sequencing with the forward and reverse M13 primers with Big Dye Terminator Reaction Mix

(Applied Biosystems). The products were purified by Big Dye XTerminator Purification Kit (Applied Biosystems) and run on ABI 3500 XL Genetic Analyzer (Applied Biosystems). Sequencing results were analyzed using the DNASTAR Lasergene 9 package. The A/New Zealand/316/2014 full genome sequence is publically available from the GISAID EpiFlu database¹ (accession numbers EPI587441-EPI587448).

Illumina MiSeq High-throughput Sequencing

The eight influenza gene segments were amplified simultaneously in a single reaction using an RT-PCR with a single set of primers (MBTuni-12 [5'-ACGCGTGATCAGCAAAGCAGG] and MBTuni-13 [5'-ACGCGTGATCAGTAGAAACAAGG]) (Zhou et al., 2009), including some modifications. Briefly, a 25 µL reaction was used for the Invitrogen Superscript III One-Step RT-PCR system with Platinum Taq DNA polymerase (Life Technologies) as per the manufacturer's instructions, including 5 µL of template RNA. The thermocycling parameters were as follows: 60 min at 42°C, 2 min at 94°C, and then 5 cycles of 30 s at 94°C, 30 s at 45°C and 3 min at 68°C, followed by 35 cycles of 30 s at 94°C, 30 s at 57°C and 3 min at 68°C, with a final extension at 68°C for 2 min. PCR amplicons were visualized by agarose gel electrophoresis, followed by purification using the QIAquick PCR purification kit (Qiagen, Valencia, CA, USA) and DNA quantitation using the Qubit dsDNA BR assay (Life Technologies).

A sequencing library was then prepared using the Illumina TruSeq DNA Nano Library preparation kit (Illumina, San Diego, CA, USA), followed by sequencing of 250-bp paired-end reads on an Illumina MiSeq instrument (New Zealand Genomics Limited, Massey Genome Service, Palmerston North, New Zealand).

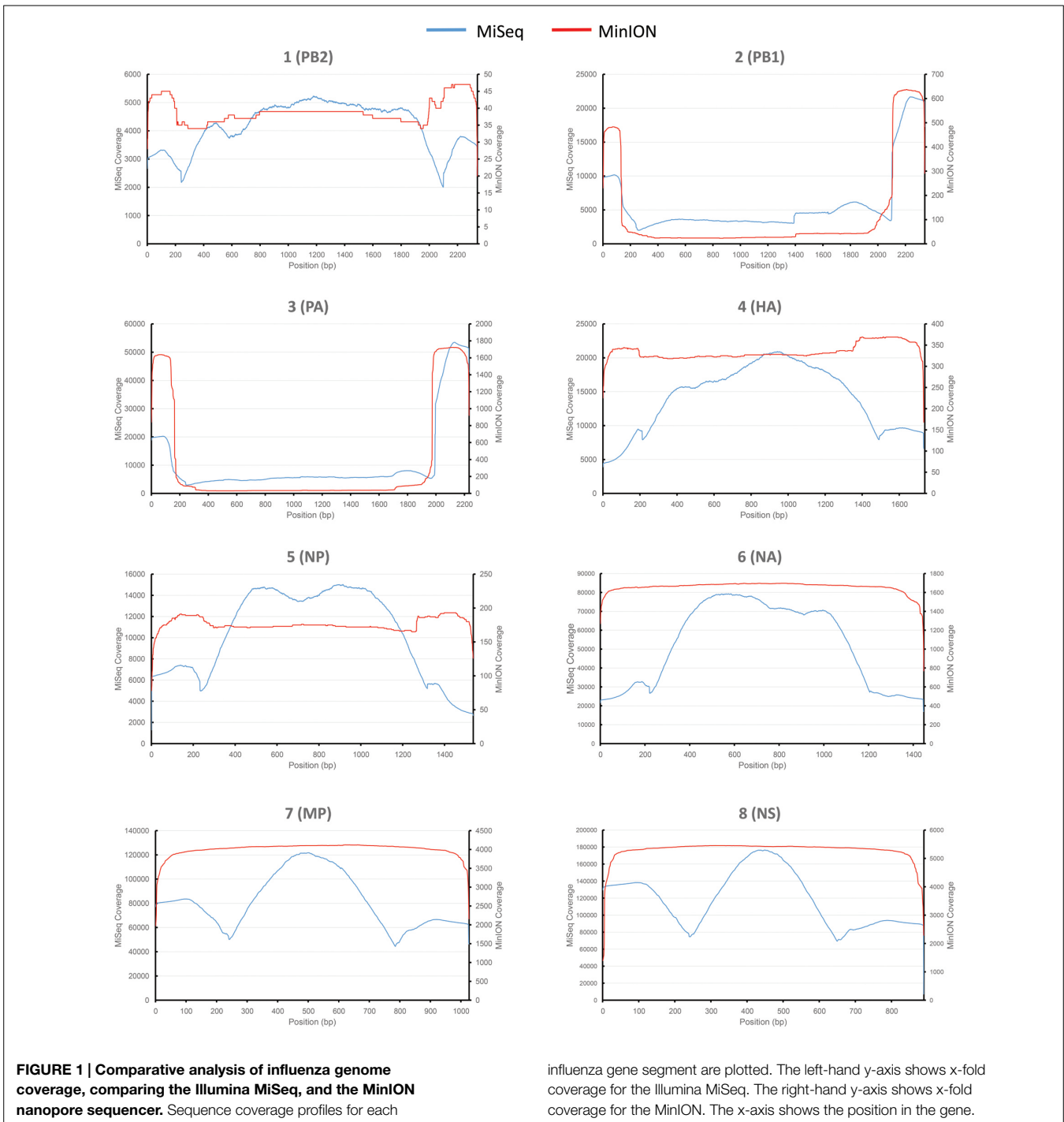
FastQC² was used to check sequence data quality. Paired-end reads were then mapped in Bowtie2 (version 2.2.4) (Langmead and Salzberg, 2012) to a reference genome Influenza A/Christchurch/503/102 (H3N2), which was accessed from the Global Initiative on Sharing Avian Influenza Data EpiFlu™ Database¹. Consensus sequence calling used SAMTools (Li et al., 2009). Data visualization was executed using Tablet (Milne et al., 2010) and Geneious R8 (Biomatters, New Zealand).

MinION Library Preparation and Sequencing

This work was completed as part of the Oxford Nanopore Technologies (ONT) MinION early-access program. The Oxford Nanopore MinION Genomic DNA Sequencing kit (version 4) was used to prepare purified PCR amplicons (as for the Illumina MiSeq; see above) for sequencing according to the manufacturer's instructions. Briefly, library preparation was as follows: 1 µg of purified PCR amplicon was diluted to 80 µL using RT-PCR molecular grade water (Ambion). The end repair process required the addition of 5 µL of Oxford Nanopore DNA CS and then use of the NEBNext End Repair module (New England Biolabs, Ipswich, MA, USA) by adding 10 µL of reaction buffer and 5 µL of enzyme, followed by a 30 min incubation at 25°C. The end repaired DNA was then purified using 100 µL Agencourt AMPure

¹<http://www.gisaid.org/>

²<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>



XP beads (Beckman Coulter Inc., Pasadena, CA, USA), as per the manufacturer's instructions with elution into 25 μ L of RT-PCR molecular grade water (Ambion). The NEBNext dA-tailing module (New England Biolabs) then required the addition of 3 μ L of 10X reaction buffer and 2 μ L of Klenow Fragment, followed by incubation of 30 min at 37°C.

8 μ L nuclease-free water was added to 30 μ L of the dA-tailed DNA, and then 10 μ L adapter mix, 2 μ L HP adapter and 50 μ L

Blunt/TA DNA ligase mastermix (New England Biolabs) were added in this order, and then incubated at 25°C for 10 min.

The ligated DNA was purified using His-tag Dynabeads. Beads were first prepared by washing in 200 μ L of 1X ONT wash buffer twice, and then resuspended in 100 μ L of 2X ONT wash buffer. The ligated DNA was mixed directly with the prepared beads and incubated at 25°C for 5 min. The beads were then washed twice again with 1X ONT wash buffer. 25 μ L of ONT elution buffer

was added to the beads with a 10 min incubation at 25°C. After removal of the beads, 6 μ L of the eluate was combined with 141 μ L EP buffer and 3 μ L fuel mix, to provide the final sequencing library.

Before loading the library, the MinION device was connected to a computer and a flow cell (R7.3) inserted. A flow cell quality control check was run using the MinKNOW control software to assess pore activity. This was followed by equilibration of the flow cell with two aliquots of 150 μ L priming mix (containing fuel mix and EP buffer), and observing a 10 min incubation between loading each aliquot.

The sequencing library was loaded into the flow cell. A 48-h sequencing protocol was selected on the MinION control software (version 0.46.1.9). The sequencing protocol was run for a total of 4 h.

MinION Data Analysis

Raw sequence data was uploaded in real-time for base calling analysis using the cloud-based Metrichor workflow rx2.22-44717-dg-1.6.1-ch-1.6.3 rev 1.9 with the quality filter on. Reads in fastq format were then extracted from downloaded HDF5 format files (fast5) using poretools (Loman and Quinlan, 2014). Mapping used the same influenza reference genome as for the Illumina MiSeq data, but was performed using LASTAL (Frith et al., 2010) with the command line: `lastal -s 2 -T 0 -Q 0 -a 1`. SAMTools (0.1.19-44428cd) was used to convert the data into bam format. Tablet (1.14.04.10) and Geneious R8 (Biomatters, New Zealand) were used for visualization.

MinION sequence data has been deposited in the European Nucleotide Archive under project accession number PRJEB9812, available at www.ebi.ac.uk/ena.

Results

430 nanopores were available on the flowcell based on the Platform QC report generated by the control software (MinKNOW) prior to sequencing run commencement. In total, 118,052 sequence reads were produced in 4 h of run time (mean length 915 bp), with 17,108 reads categorized as high quality two-direction (2D) reads (mean length 955 bp). Sequence alignment to a prototypic reference genome was undertaken using the software package LASTAL, and full coverage of all influenza genes was achieved (Figure 1). Influenza sequence accounted for over 86.46% (14,793 reads) of the available 2D sequence data (Table 1).

For comparison, single-tube PCR amplicons from the same influenza A (H3N2) influenza isolate were run on an Illumina MiSeq, producing 2,999,678 sequence reads of 250 bp in length. Bowtie2 was used to map sequence data to the same prototypic reference genome, resulting in 100% coverage, just as observed for the MinION nanopore sequencer (Figure 1). The coverage for the MiSeq was higher than for the MinION, due to the higher number of reads generated in the Illumina MiSeq run. The coverage for the MinION appears to be more even across the entire gene segment for HA, NA, NP, MP, and NS, as compared to coverage observed for the MiSeq.

The Illumina MiSeq dataset showed 100% identity with the Sanger-sequence genome of the influenza (H3N2) virus

(Figure 2). The MinION nanopore sequencer also showed high concordance with the Sanger-sequence data, having greater than 99% pairwise identity across the genome (Figure 2).

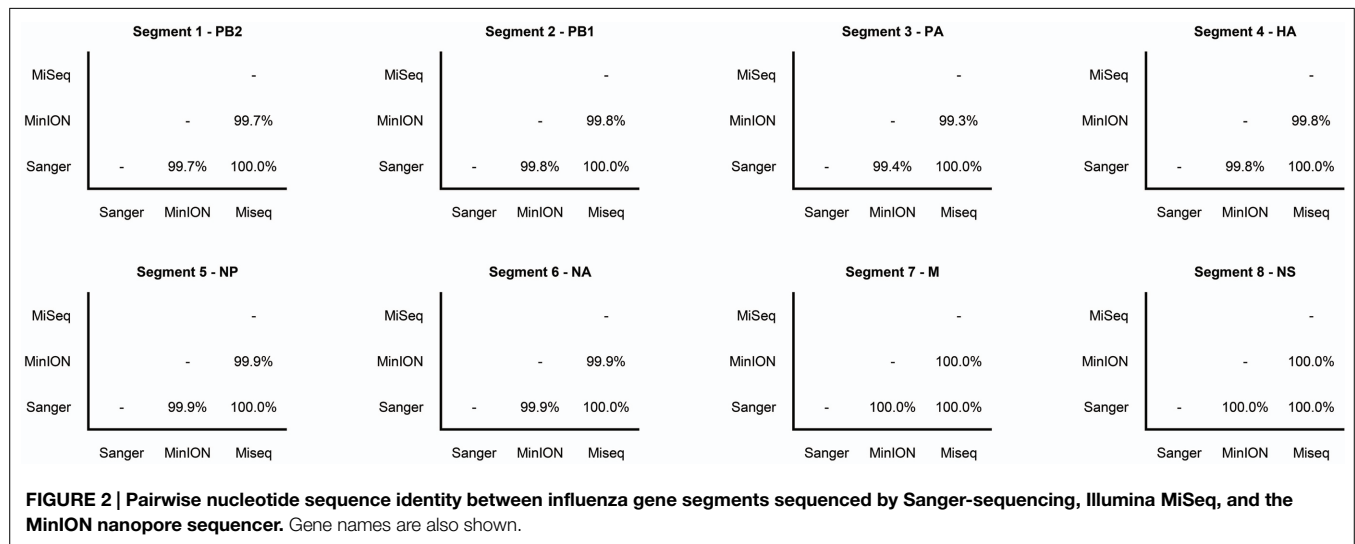
Discussion

A whole genome sequence of an influenza virus was produced on the MinION nanopore sequencer using amplicon sequencing. All eight influenza gene segments from an influenza A (H3N2) virus were amplified simultaneously in a single PCR tube. This PCR method has previously been published, where primers target the conserved 12/13 base-pair viral RNA termini common to each influenza gene, and are conserved across all influenza A subtypes (Zhou et al., 2009). Amplicons were placed through the standard genomic DNA MinION workflow. The eight gene segments range in size from 0.89 to 2.3 kilobases, and amplification bias was apparent as reflected in the differing coverage for each gene segment (Figure 1) — an effect that has previously been documented (Van den Hoecke et al., 2015). Refined methods have now become available to limit this bias, which include adjustments of PCR primer concentrations, annealing temperature and elongation time (Van den Hoecke et al., 2015). The nucleotide discrepancies observed between the genome generated by the MinION as compared to the other methods are minimal but will still need to be resolved before the technique can be applied more widely. These discrepancies were distributed randomly throughout the genome and are most likely the result of sequencing error, as has been documented by others (Mikheyev and Tin, 2014; Quick et al., 2014; Kilianski et al., 2015). The per base accuracy of the MinION has been reported at between 65 and 80% (Kilianski et al., 2015). Therefore, discrepant nucleotides in the MinION-generated genome are not likely to represent genuine mutations that were missed by either Sanger-sequencing or the Illumina MiSeq. Point mutations can have significant phenotypic effects on influenza viruses, for example, in the neuraminidase gene they may confer resistance to antiviral drugs such as oseltamivir (Yen et al., 2006). It is therefore important that 100% identity with conventional sequencing approaches is attained before deployment in the clinic or field.

It is also unclear why the a proportion of the sequence reads generated in the MinION dataset are longer than the gene segment length, i.e. gene segment 4 (haemagglutinin) is 1,733 bp in length but records a maximum sequence read of 2,766 bp in the alignment (Table 1). This could be accounted for by the generation of chimeric amplicons during the PCR, where incomplete extension during the elongation phase of PCR provides heterologous templates for primers to bind to (Smyth et al., 2010). We could not provide direct evidence for this occurring but such an effect would not be able to be detected by a short-read NGS technology, even in the event where a chimeric junction is sequenced, as constraints of the mapping process would exclude such short reads from the final assembly. Similarly, Sanger-sequencing would be unlikely to reveal such chimera if they were rare enough (for example < 5%). However, an alternative explanation must be considered that an artifact of analysis or detection in the MinION instrument accounts for the apparent longer read lengths. Mapping of the MinION data

TABLE 1 | Sequence read lengths and the proportions of reads assigned to the influenza genome.

Gene segment (Gene name)	1 (PB2)	2 (PB1)	3 (PA)	4 (HA)	5 (NP)	6 (NA)	7 (M)	8 (NS)	Total
Gene length (bp)	2,341	2,341	2,233	1,733	1,537	1,448	1,027	890	13,550
No of 2D reads assigned	59	702	1,931	395	216	1,755	4,191	5,544	14,793
Percentage of total 2D reads	0.34%	4.10%	11.29%	2.31%	1.26%	10.26%	24.49%	32.41%	86.46%
Mean read length (bp)	1,243	245	490	1,498	1,259	1,418	1,045	913	–
Median read length (bp)	1,961	448	465	1,765	1,519	1,472	1,078	927	–
5th percentile (bp)	452	376	394	563	464	776	676	711	–
95th percentile (bp)	2,406	1,088	898	1,878	1,655	1,559	1,134	984	–
Minimum read length (bp)	313	206	211	369	343	223	171	283	–
Maximum read length (bp)	2,473	2,351	2,347	2,766	2,045	2,366	2,045	2,766	–



using BWA (Li, 2013) was found to provide very similar results to LAST (data not shown), indicating that the mapping software is probably not the cause of such artifacts. *In silico* generation of chimeric sequence could occur if strand displacement at the nanopore is not recognized by the instrument during a run. This is speculation and further work will be required to resolve the reason for the increased apparent read lengths relative to the size of the original gene segments (data available from the European Nucleotide Archive; accession number PRJEB9812). It must be noted that the majority of MinION reads are of comparable length to their respective gene segments, i.e. 95% of reads mapping to HA are less than 1,878 bp, compared to the HA gene length of 1,733 bp (Table 1). The median read length for HA, NA, M, and NS is slightly longer than the gene segment length (Table 1), which could be due to sequencing of adaptors or insertions caused by sequencing error. Blast search of the unmapped portion of sequences failed to show any similarity with known sequences in genbank (data not shown).

The PA and PB1 genes appear to have higher coverage at the 5' and 3' termini, but lower coverage in the middle of the gene (Figure 1). This effect is observed in both MiSeq and MinION data. It is apparent to a lesser extent in the PB2 gene (probably masked by the low coverage in this gene). This is likely to have been caused by the presence of defective interfering viral particles,

and very similar coverage profiles for NGS data have been noted before for PB2, PB1, and PA genes (Saira et al., 2013). Defective interfering virus particles of influenza have incomplete genomes, especially having truncated copies of PB2, PB1 and PA genes (Saira et al., 2013). This has resulted in an apparent lower median (and mean) read length for these genes in the MinION data (Table 1), which is an artifact of the short amplicons generated by the defective particles, rather than from the sequencing process itself.

The base-calling analysis for the MinION nanopore sequencer is an area of active development, as is the development of flowcells and associated enzymatic chemistries. The best quality sequence data are termed 2D reads. These are the consensus reads produced by comparing both the template and complementary strand sequences that are obtained when the double-helix is unwound and passed through the nanopore as a single-strand (Quick et al., 2014). Currently there are two technical aspects which affect the ability to produce 2D reads, those being the accuracy of measuring the electropotential across the nanopore, and the efficiency of the motor protein and hairpin adaptor binding. In the present study, 2D data account for 14.49% of the total reads. This figure is lower than previously published work by Quick et al (Quick et al., 2014) who report 22.6% 2D data when using the R7.3 flowcell. Although the accuracy of MinION sequence data is an area of active

development and increases in data quality are anticipated (Bayley, 2015), this study shows the platform is already suitable for rapid influenza virus detection and provision of a draft genome. A 4 h run time for the MinION provides enough data to generate a reliable genome. But in order to attain 100% identity with other methods it will be necessary to increase sequence depth, most likely by an increase in run time.

Other single molecule sequencers, i.e., Helicos (Thompson and Steinmann, 2010) and Pacific Biosciences (Flusberg et al., 2010; McCarthy, 2010), require significant amounts of laboratory space, specialized equipment and large-scale computing resources. The MinION sequencer is small and portable. Run control and data analysis in this study were done in real-time using a moderate specification laptop (minimum hardware requirements: i7 CPU, 8 Gb RAM, and 128 GB solid state hard disk). This is in direct contrast to the high-end computing resources and expertise required to process short-read high throughput sequencing data for other platforms. The small footprint, potential for portability, and the use of common molecular biology equipment and reagents are attractive factors which show the MinION nanopore sequencer has potential in diagnostic applications. Furthermore, the short-run time of 4 h utilized in this study is a realistic timeframe for the production of clinically useful information. This

study shows it is conceivable that investigators working in the field, for example, during an influenza pandemic or outbreak, could analyze samples on-site and draw initial conclusions on subtype, reassortants, pathogenicity and antiviral drug sensitivity (Barzon et al., 2013; Jones, 2015). The MinION nanopore sequencer can provide a useful method for diagnostic, research and reference laboratories to readily obtain full length influenza virus genomes.

Acknowledgments

This study was funded by the ESR Core Research Fund from the New Zealand Ministry of Business, Innovation and Employment. The Melbourne WHO Collaborating Centre for Reference and Research on Influenza is supported by the Australian Government Department of Health. We thank Sue Huang and staff in ESR Clinical Virology and for their support and provision of the isolate. We also acknowledge NZGL and the Massey Genome Service for their high-throughput sequencing services, especially Patrick Biggs and Lorraine Berry for their advice and guidance. We also gratefully acknowledge the originating laboratory, Canterbury Health Services (New Zealand), who provided the reference influenza genome; available on the GISAID EpiFlu™ Database.

References

- Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., et al. (2015). MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat. Biotechnol.* 33, 296–300. doi: 10.1038/nbt.3103
- Barzon, L., Lavezzo, E., Costanzi, G., Franchin, E., Toppo, S., and Palu, G. (2013). Next-generation sequencing technologies in diagnostic virology. *J. Clin. Virol.* 58, 346–350. doi: 10.1016/j.jcv.2013.03.003
- Bayley, H. (2015). Nanopore sequencing: from imagination to reality. *Clin. Chem.* 61, 25–31. doi: 10.1373/clinchem.2014.223016
- Bertelli, C., and Greub, G. (2013). Rapid bacterial genome sequencing: methods and applications in clinical microbiology. *Clin. Microbiol. Infect.* 19, 803–813. doi: 10.1111/1469-0691.12217
- Deng, Y.-M., Spirason, N., Iannello, P., Jelley, L., Lau, H., and Barr, I. G. (2015). A simplified Sanger sequencing method for full genome sequencing of multiple subtypes of human influenza A viruses. *J. Clin. Virol.* 68, 43–48. doi: 10.1016/j.jcv.2015.04.019
- Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., et al. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* 7, 461–465. doi: 10.1038/nmeth.1459
- Frith, M. C., Hamada, M., and Horton, P. (2010). Parameters for accurate genome alignment. *BMC Bioinformatics* 11:80. doi: 10.1186/1471-2105-11-80
- Ghedini, E., Sengamalay, N. A., Shumway, M., Zaborosky, J., Feldblyum, T., Subbu, V., et al. (2005). Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* 437, 1162–1166. doi: 10.1038/nature04239
- Hall, R. J., Wang, J., Todd, A. K., Bissielo, A. B., Yen, S., Strydom, H., et al. (2014). Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J. Virol. Methods* 195, 194–204. doi: 10.1016/j.jviromet.2013.08.035
- Jones, B. (2015). Technology: nanopore sequencing for clinical diagnostics. *Nat. Rev. Genet.* 16, 68. doi: 10.1038/nrg3895
- Kilianski, A., Haas, J. L., Corriveau, E. J., Liem, A. T., Willis, K. L., Kadavy, D. R., et al. (2015). Bacterial and viral identification and differentiation by amplicon sequencing on the MinION nanopore sequencer. *Gigascience* 4, 12. doi: 10.1186/s13742-015-0051-z
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at: arXiv:1303.3997v2 [q-bio.GN].
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, L., Deng, X., Mee, E. T., Collet-Teixeira, S., Anderson, R., Schepelmann, S., et al. (2015). Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. *J. Virol. Methods* 213, 139–146. doi: 10.1016/j.jviromet.2014.12.002
- Loman, N. J., and Quinlan, A. R. (2014). Poretools: a toolkit for analyzing nanopore sequence data. *Bioinformatics* 30, 3399–3401. doi: 10.1093/bioinformatics/btu555
- Madoui, M. A., Engelen, S., Cruaud, C., Belser, C., Bertrand, L., Alberti, A., et al. (2015). Genome assembly using Nanopore-guided long and error-free DNA reads. *BMC Genom.* 16:327. doi: 10.1186/s12864-015-1519-z
- McCarthy, A. (2010). Third generation DNA sequencing: pacific biosciences single molecule real time technology. *Chem. Biol.* 17, 675–676. doi: 10.1016/j.chembiol.2010.07.004
- McGinn, S., and Gut, I. G. (2013). DNA sequencing—spanning the generations. *N. Biotechnol.* 30, 366–372. doi: 10.1016/j.nbt.2012.11.012
- Mikheyev, A. S., and Tin, M. M. (2014). A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.* 14, 1097–1102. doi: 10.1111/1755-0998.12324
- Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., et al. (2010). Tablet—next generation sequence assembly visualization. *Bioinformatics* 26, 401–402. doi: 10.1093/bioinformatics/btp666
- Nakamura, S., Yang, C. S., Sakon, N., Ueda, M., Tougan, T., Yamashita, A., et al. (2009). Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *PLoS ONE* 4:e4219. doi: 10.1371/journal.pone.0004219
- Quick, J., Quinlan, A. R., and Loman, N. J. (2014). A reference bacterial genome dataset generated on the MinION portable single-molecule nanopore sequencer. *Gigascience* 3, 22. doi: 10.1186/2047-217X-3-22
- Quinones-Mateu, M. E., Avila, S., Reyes-Teran, G., and Martinez, M. A. (2014). Deep sequencing: becoming a critical tool in clinical virology. *J. Clin. Virol.* 61, 9–19. doi: 10.1016/j.jcv.2014.06.013
- Saira, K., Lin, X., Depasse, J. V., Halpin, R., Twaddle, A., Stockwell, T., et al. (2013). Sequence analysis of *in vivo* defective interfering-like RNA of influenza A H1N1 pandemic virus. *J. Virol.* 87, 8064–8074. doi: 10.1128/jvi.00240-13

- Smyth, R. P., Schlub, T. E., Grimm, A., Venturi, V., Chopra, A., Mallal, S., et al. (2010). Reducing chimera formation during PCR amplification to ensure accurate genotyping. *Gene* 469, 45–51. doi: 10.1016/j.gene.2010.08.009
- Thompson, J. F., and Steinmann, K. E. (2010). Single molecule sequencing with a HeliScope genetic analysis system. *Curr. Protoc. Mol. Biol.* Chapter 7:Unit7 10. doi: 10.1002/0471142727.mb0710s92
- Van den Hoecke, S., Verhelst, J., Vuylsteke, M., and Saelens, X. (2015). Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing. *BMC Genom.* 16:79. doi: 10.1186/s12864-015-1284-z
- Vernikos, G., Medini, D., Riley, D. R., and Tettelin, H. (2015). Ten years of pan-genome analyses. *Curr. Opin. Microbiol.* 23, 148–154. doi: 10.1016/j.mib.2014.11.016
- Yen, H. L., Hoffmann, E., Taylor, G., Scholtissek, C., Monto, A. S., Webster, R. G., et al. (2006). Importance of neuraminidase active-site residues to the neuraminidase inhibitor resistance of influenza viruses. *J. Virol.* 80, 8787–8795. doi: 10.1128/jvi.00477-06
- Zhou, B., Donnelly, M. E., Scholes, D. T., St George, K., Hatta, M., Kawaoka, Y., et al. (2009). Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza A viruses. *J. Virol.* 83, 10309–10313. doi: 10.1128/jvi.01109-09

Conflict of Interest Statement: This work was done as part of the MinION Access Programme (MAP). The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Wang, Moore, Deng, Eccles and Hall. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.