Nanopore-based DNA sequencing in clinical microbiology: preliminary

assessment of basic requirements

Håvard Harstad^a, Rafi Ahmad^{b*}, Anders Bredberg^{a*}

^a*Department of Medical Microbiology, Innlandet Hospital Trust, Lillehammer, Norway*

^b*Department of Natural Sciences and Technology, Faculty of Education and Natural Sciences, Inland Norway University of Applied Sciences, Hamar, Norway.*

*Corresponding authors. E-mail addresses: anders.bredberg@sykehuset-innlandet.no
rafi.ahmad@inn.no

Keywords: Nanopore sequencing, metagenomic analysis, clinical diagnostics

Abstract

Aim: Identify basic requirements for a metagenomic nanopore sequencing protocol

permitting frequent application in a clinical microbiology daily routine diagnostic setting.

Background: Nanopore sequencing with the Oxford Nanopore Technologies MinION device has a potential to markedly improve clinical diagnosis of infections. Reports have emerged recently that it may provide direct-from-clinical-sample information; for example, with urine samples, bronchial tuberculosis samples and orthopedic prostheses. However, the ideal protocol for clinical use remains to be determined, especially relating to detection of relevant pathogen quantities and to finding a reasonable level of economic costs.

Results: MinION can provide qualitatively and quantitatively correct identification of multiple species in metagenomics samples. For detection of clinically relevant quantities of bacteria (on a nanogram DNA level) there is a need for carrier DNA. Importantly, high-purity DNA and a naïve MinION flow cell seem to be critical parameters.

Conclusions: Our results suggest that high-purity clinical sample DNA, addition of carrier DNA and a naïve flow cell are critical factors for clinical use of MinION. A relatively high error rate may limit detection of antimicrobial resistance genes, and a realistic level of costs will require availability of a price-reduced and single-use flowcell.

Introduction

Nanopore-based metagenomic sequencing with the MinION device is a promising new tool for clinical diagnostic work (1-2). With the nanopore technique, you read the bases in a DNA strand as it threads through a nanometer-sized pore, by detecting changes in an ionic current across the pore as DNA bases pass through. Because it does not require DNA to be chopped into small pieces, as most other sequencing technologies do, you can get longer reads, which can be assembled more quickly into a full sequence. Another advantage over other metagenomics techniques is the real-time mode of analysis. For both RNA and DNA pathogens, the sequencer produces an immediate readout of the strand going through the pore (3).

This new technology will most probably become a standard part of routine clinical work for rapid identification and characterization of microbial isolates. The great potential of the technique has been described in Science: ‘it will make a tremendous difference both to diagnostics and to science it will quickly make its way into many labs and the clinic’ (4). As illustrations of its clinical usefulness in threatening situations, a hemorrhagic fever virus (Ebola) was identified within 6 hours (5), and same-day diagnostic results for tuberculosis has been reported with the nanopore MinION device (6). The method has also been used for identification of viruses in a clinical sample (7), for bacterial pathogens in urine (8), in orthopedic prostheses-related tissue (9), in a liver abscess (10), in faeces (11) and to analyse for the anthrax agent (12) and to predict antibiotic resistance pattern in gonorrhea (13). Nevertheless, this clinical research field is quite new, and, as yet, the available studies have been on limited groups of pathogens and on few samples.

The aim of this pilot project was limited to exploring the possibility to use nanopore metagenomic sequencing for cost-beneficial, rapid and at informative diagnostic work in clinical microbiology.

Methods

Microbiology

The following reference bacterial strains were obtained from the American Type Culture Collection (ATCC): *Escherichia coli* ATCC 25922, *Klebsiella pneumoniae* ATCC 700603 expressing extended spectrum beta lactamase (ESBL) resistance (14), *Pseudomonas aeruginosa* ATCC 27853, *Proteus mirabilis* ATCC 29906 and *Haemophilus influenza* ATCC 49766. *Acinetobacter baumannii* is an in-house reference strain, Matrix Assisted Laser Desorption Ionization Time of Flight (MALDI-TOF MS microflex LT from Bruker Daltonics, Bremen, Germany) was used, with scores >2 verifying the strain identity before analysis.

DNA extraction and sequencing

DNA was extracted with a robot employed by our clinical diagnostic laboratory for PCR analysis of patient samples (Roche MagnaPure Compact), or manually with standard silica mini-columns (Qiagen Genomic-tip 20/G) following the manufacturer's instruction. DNA purity and concentration were determined using Nanodrop (NanoDrop 2000, ThermoFisher Scientific). Approximately 400-500 ng DNA was taken to build a DNA library for nanopore sequencing using a Rapid Sequencing Kit (SQK-RAD003 from Oxford Nanopore Technologies, Oxford, UK (ONT)) as described by the manufacturer, and then loaded onto a MinION sequencing apparatus flow cell. The MINION sequencing was performed with R9.4 flow cells (FLO-MIN106, ONT).

Bioinformatics analysis of the obtained metagenomic sequences

The MinION sequencing was controlled using ONT's MinKNOW software. The version of the software used varied from run to run but can be determined by inspection of fast5 files from the data set. Raw sequence reads were basecalled in real time via MinKNOW producing .fastq files. Basecalled data were then uploaded to the Epi2Me interface, a platform for cloud-based analysis of MinION data. Two different applications in Epi2Me were applied for data interpretation, What's In My Pot? (WIMP) for identification of sequence data, and Antimicrobial Resistance Application (ARMA) software for identification of resistance genes.

In a number of instances, MinION sequencing control was shifted to customized MinKNOW scripts. These scripts provided enhanced pore utilization/data yields during sequencing, and operated by monitoring and adjusting flow cell bias-voltage (-180 mV to -250 mV). More detailed information on these scripts can be found on the ONT community website. It is also worth noting that both software applications and chemistry in reagent kits and flow cells are under continuous development at ONT and updates appear frequently. As an example, the sequencing kit used in this report, SQK-RAD003, has been replaced by SQK-RAD004 having an updated composition. Furthermore, software updates often tends to imply changes in the interface + resetting all run scripts (if edited).

Results and Discussion

Our first experiments followed a recently published clinical protocol for urinary tract infections (8). A urine sample submitted to our clinical laboratory and reported to the submitting clinician to have a clinically significant quantity ($>10^8/\text{L}$) of *P. mirabilis*, as determined by our current standard diagnostic methods with Petri dish agar medium culturing

and MALDI-TOF species identification, was used. After the sample had been anonymized, DNA was extracted on the Magnapure Compact in accordance with the published protocol. The purity and yield of the extracted DNA did not meet the requirements stipulated by the library SQK-RAD003 kit, but nevertheless a library was made. MinION sequencing showed dominance (77%) of eukaryotic (and presumably human) reads and 10% *P. mirabilis* reads. Furthermore the experiment produced very little data: 650 kb were produced within 2.5 hr. Because the results obtained by us with this protocol did not correspond to our needs for bacterial quantity sensitivity or DNA purity, and because of the low fraction of bacterial reads, we decided to focus on more basic protocol components involving DNA quality, MinION device and ONT software. All following experiments were performed without clinical samples, and instead with defined counts of bacterial reference strains, suspended in saline, and presumed problems associated with human samples such as DNA-degrading inflammatory substances and large quantities of patient DNA (15) were postponed until a later developmental phase.

Bacterial cultures at an 0.5 McFarland density, corresponding to approximately 10^8 bacterial cells/ml, of *E. coli* and *K. pneumoniae*, respectively, were mixed and DNA was extracted using both the MagnaPure Compact robot and manually with Qiagen Genomic-tip 20/G mini-columns. The results clearly show that the mini-column extraction is superior; there were more of informative reads and greater fragment lengths, obtained within a shorter real-time after loading onto the MinION flow cell (Table 1). For *E. coli* the fraction of informative reads was 48% and 92%, with Magnapure and mini columns, respectively; and the corresponding values for *K. pneumoniae* were 14% and 92%. Another observation was the longer read fragments with the mini-columns, likely to be due to gentle handling of the DNA during extraction.

Table 1. Sequencing results for a mixed *E. coli* and *K. pneumoniae* sample, with two different DNA isolation protocols.

Microorganism	DNA isolation method	MinION run time (hr)	Reads, informative (total)	Total yield (Mbp)	Best hit (reads)	Average length (kb)
<i>E. coli</i>	MagnaPure Compact	6	993 (2082)	9.6	<i>E. coli</i> (350)	4.6
<i>E. coli</i>	Qiagen Mini-column	2	25847 (28191)	326.3	<i>E. coli</i> (24654)	11.6
<i>K. pneumoniae</i>	MagnaPure Compact	48	914 (6602)	15.3	<i>K. pn</i> (117)	2.3
<i>K. pneumoniae</i>	Qiagen Mini-column	24	43953 (48021)	541.7	<i>K. pn</i> (39082)	11.3

Detection of antimicrobial resistance genes was carried out on selected data by using the Antimicrobial resistance application (AMRA). The *K. pneumoniae* mini-column sequence data described in Table 1 was subjected to further investigation by uploading on AMRA via the cloud-based Epi2Me agent. The strain is known to contain resistance genes, including the beta lactamase OKP-B-7 gene and the plasmid pKQPS2 resistance genes aadB (aminoglycoside), sul1 (sulfonamide), and ESBL encoding genes OXA-2 and SHV-18 (14). The report contained limited information, with 120 resistance gene hits, with various confidences and bacterial species. For *K. pneumoniae* AMRA only reported OKP-B-7 and SHV-18. These results apparently reflect the relatively high error rate (10-20%) in MinION sequencing reads. The report is available when logging in to Methrichor.

Next, we decided to attempt co-detection of 5 bacterial pathogens, within one sample, at relatively low concentrations ranging from as little as 5 ng DNA and downwards in a 10-log series, and also to test a re-used MinION flow cell. Because the MinION SQK-RAD003 protocol requires an input of at least 400 ng DNA, corresponding to a quantity as high as approximately 10^8 bacterial cells, and because we are seeking a direct-from-clinical-sample protocol with no DNA amplification step, we furthermore decided to add each low concentration sample to 400 ng DNA from the low-pathogen *A. baumannii* to function as

carrier. The DNA sequencing of this metagenomic sample did not give satisfactory results. During a 7 hr run, only 6280 reads were generated with a total yield of 102.6 Mb. None of the low concentration microbes were identified by WIMP, only *A. baumannii* (5070 reads). A sudden drop in available pores in the flow cell was observed as the experiment started, from 1003 down to 251 available pores. A separate experiment with a mixture of *A. baumannii* and *P. aeruginosa* provided similar disappointing results. Because our previous experiment related in Table 1 was performed with a naïve flow cell, the results were suggestive to us that flow cell reuse is associated with a significant and unacceptable drop in sensitivity and data accumulation even with DNA isolated on mini-columns.

Next, using manually mini-column extracted DNA from 4 different bacteria and using a naïve flow cell, we wanted to analyze a truly metagenomic sample. A high quantity (292 ng) of *P. mirabilis* DNA served as carrier, and was mixed with lower amounts of *H. influenzae* (105 ng), *K. pneumoniae* (52 ng) and *E. coli* (10 ng corresponding to approximately 2×10^6 bacterial cells). A SQK-RAD003 sequencing library was constructed from the pool of DNA showing high purity (OD 260/280 ~1.8 and OD 260/230 2.0-2.2).

Figure 1. DNA sequencing results from a metagenomic sample containing *P. mirabilis*, *H. influenzae*, *K. pneumoniae* and *E. coli*.

All four species were accurately identified (Figure 1), and the read numbers corresponded well with the genome content for each bacterium in the sequencing library; indicating that MinION can both qualitatively and quantitatively provide correct metagenomic data. Furthermore, there was a trend towards a relatively high read number as compared with colony count at the lower quantity range, suggesting a MinION sensitivity approaching clinically relevant concentrations at the ng level (corresponding to approximately 10^5 bacterial cells); in agreement with the data reported recently on orthopedic device infection samples (9). These results indicate that a gentle manual DNA extraction procedure and single-use flow cells are critical parameters.

Our combined preliminary assessment of MinION suggests to us that it readily identifies multiple microorganisms in a metagenomic manner. First and foremost MinION promises to become able to function as a rapid and accurate pathogen species identification technique for clinical specimens. The technology is under continued development, and refinement of sample preparation will most probably pose a major challenge, with a need to customize it according to sample type.

Conflict of interest statement

All the authors declare no conflict of interests.

References

1. Bailey H. Nanopore sequencing; from imagination to reality. *Clin Chem* 61, 25-31. 2015.
2. Judge K, Hunt M, Teuter S, Tracey A, Quail MA, Parkhill J, Peacock SJ. Comparison of bacterial genome assembly software for MinION data and their applicability to medical microbiology. *Microb Genom* 8, e000085, 2016.
3. McCabe MS, Cormican P, Johnston D, Earley B. Simultaneous detection of DNA and RNA virus species involved in bovine respiratory disease by PCR-free rapid tagmentation-

- based library preparation and MinION nanopore sequencing. *BioRxiv* 2018.
<http://dx.doi.org/10.1101/269936>
4. Pennisi E. Pocket-sized sequencers start to pay off big. *Science* 356 (6338), 572-573, 2017.
5. Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med* 7, 99-, 2015.
6. Votintseva AA, Bradley P, Pankhurst L, del Ojo Elias C, Loose M, Nilgiriwala K, et al. Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J Clin Microbiol* 55, 1285-1298, 2017.
7. Gunther T, Haas L, Alawi M, Wohlsein P, Marks J, Grundhoff A, et al. Recovery of the first full-length genome sequence of a parapoxvirus directly from a clinical sample. *Sci Rep* 7:3734, 2017.
8. Schmidt K, Mwaigwisya S, Crossman LC, Doumith M, Munroe D, Pires C, et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J Antimicrob Chemother* 72, 104-114, 2017.
9. Sanderson ND, Street TL, Foster D, Swann J, Atkins BL, Brent AJ. Real-time analysis of nanopore-based metagenomic sequencing from orthopaedic device infection. *BioRxiv* 2017.
<http://dx.doi.org/10.1101/220616>
10. Gong L, Huang Y-T, Wong C-H, Chao W-C, Wu Z-Y, Wei C-L, Liu P-Y. Culture-independent analysis of liver abscess using nanopore sequencing. *PLoS ONE* 13, e0190853, 2018.
11. Leggett RM, Alcon-Giner C, Heavens D, Caim S, Brook TC, Kujawska M, et al. Rapid MinION metagenomic profiling of the preterm infant gut microbiota to aid in pathogen diagnostics. *BioRxiv* 2017. <http://dx.doi.org/10.1101/180406>
12. Tyler AD, Mataseje L, Urfano CJ, Schmidt L, Antonation KS, Mulvey MR, Corbett CR. Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. *Sci Rep* 8, 10931, 2018.
13. Phillips LT, Witney A, Izquierdo-Carrasco F, Mayes S, Wright A, Laing K, et al. Hand-held rapid whole genome nanopore sequencing to predict neisseria gonorrhoeae antibiotic susceptibility: steps towards clinic based tailored antimicrobial therapy. *Sex Transm Infect* 93 (Suppl 2), A56, 2017.
14. Elliott AG, Ganesamoorthy D, Coin L, Cooper MA, Duc Cao M. Complete genome sequence of *Klebsiella quasipneumoniae* supsp. *similipneumoniae* strain ATCC 700603. *Genome Announc* 4, e00438, 2016.
15. Handschur M, Karlic H, Hertel C, Pfeilstocker M, Haslberger AG. Preanalytic removal of human DNA eliminates false signals in general 16S rDNA PCR monitoring of bacterial pathogens in blood. *Comp Immunol Microb Inf Dis* 32, 207-219, 2009.