

RESEARCH ARTICLE

Field-based detection of bacteria using nanopore sequencing: Method evaluation for biothreat detection in complex samples

Andrea D. Tyler¹, Jane McAllister², Helen Stapleton³, Penny Gauci², Kym Antonation¹, David Thirkettle-Watts², Cindi R. Corbett^{1,4*}

1 National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada, **2** Department of Defence, Defence Science and Technology Group, Fishermans Bend, Melbourne, VIC, Australia, **3** Defence Science and Technology Laboratory, Dstl Porton Down, Salisbury, Wiltshire, United Kingdom, **4** Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Winnipeg, Manitoba, Canada

* cindi.corbett@canada.ca



OPEN ACCESS

Citation: Tyler AD, McAllister J, Stapleton H, Gauci P, Antonation K, Thirkettle-Watts D, et al. (2023) Field-based detection of bacteria using nanopore sequencing: Method evaluation for biothreat detection in complex samples. *PLoS ONE* 18(11): e0295028. <https://doi.org/10.1371/journal.pone.0295028>

Editor: Farah Al-Marzooq, UAE University: United Arab Emirates University, UNITED ARAB EMIRATES

Received: May 10, 2023

Accepted: November 14, 2023

Published: November 28, 2023

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0295028>

Copyright: © 2023 Tyler et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Samples available as part of NCBI BioProject: PRJNA943266.

Abstract

From pathogen detection to genome or plasmid closure, the utility of the Oxford Nanopore Technologies (ONT) MinION for microbiological analysis has been well documented. The MinION's small footprint, portability, and real-time analytic capability situates it well to address challenges in the field of unbiased pathogen detection, as a component of a security investigation. To this end, a multicenter evaluation of the effect of alternative analytical approaches on the outcome of MinION-based sequencing, using a set of well-characterized samples, was explored in a field-based scenario. Three expert scientific response groups evaluated known bacterial DNA extracts as part of an international first responder (Chemical, Biological, Radiological) training exercise. Samples were prepared independently for analysis using the Rapid and/or Rapid PCR sequencing kits as per the best practices of each of the participating groups. Analyses of sequence data were in turn conducted using varied approaches including ONTs What's in my pot (WIMP) architecture and in-house computational pipelines. Microbial community composition and the ability of each approach to detect pathogens was compared. Each group demonstrated the ability to detect all species present in samples, although several organisms were detected at levels much lower than expected with some organisms even falling below 1% abundance. Several 'contaminant' near neighbor species were also detected, at low abundance. Regardless of the sequencing approach chosen, the observed composition of the bacterial communities diverged from the input composition in each of the analyses, although sequencing conducted using the rapid kit produced the least distortion when compared to PCR-based library preparation methods. One of the participating groups generated drastically lower sequencing output than the other groups, likely attributed to the limited computer hard drive capacity, and occasional disruption of the internet connection. These results provide further consideration for conducting unbiased pathogen identification within a field setting using MinION sequencing. However, the benefits of this approach in providing rapid results and unbiased detection must be considered along with the complexity of sample preparation

Funding: This work was supported by the Canadian Safety and Security Program project, CSSP-2018-TI-2372. CR Corbett, KS Antonation were recipients of the CSSP funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Note the opinions expressed within do not represent the opinions of the Public Health Agency of Canada or the Government of Canada.

Competing interests: The authors have declared that no competing interests exist.

and data analytics, when compared to more traditional methods. When utilized by trained scientific experts, with appropriate computational resources, the MinION sequencing device is a useful tool for field-based pathogen detection in mixed samples.

Introduction

In the rapidly evolving field of bioforensics and biothreat response, application of novel methods to analyze various samples in non-laboratory settings is critical. The capability to detect a wide variety of pathogens without *a priori* knowledge of the potential threat agent in a non-laboratory setting is the ideal. As such, Oxford Nanopore Technologies (ONT) MinION device, which can rapidly produce long read sequence data for agnostic biothreat detection and characterization, is ideally situated: it is small and highly portable, generates long read sequence data, has the ability to analyze reads almost immediately following the start of a sequencing run, and can be used for a variety of sample types, including mixed environmental samples [1–3]. Previous work has demonstrated this technologies' utility in the context of field-based detection of Ebola, Zika, and tuberculosis (TB), and it has been previously shown to have use in remote settings simulating biothreat response [4–6]. Despite this body of work and several sources describing the strengths and limitations of this method [7–9], none to date have documented the capability of the MinION device on a set of well characterized samples in the field, as part of a blinded, multicentre assessment. To this end, teams from Australia, Canada and the United Kingdom evaluated the capability of the MinION sequencing platform for identification of bacterial agents of interest in a set of mixed DNA samples as part of an international first responder exercise.

Materials and methods

Prior to the exercise, sample DNA was prepared by Group 1 as follows: Bacterial strains were incubated at 37°C overnight in LB broth, and DNA was extracted using the Blood and Cell Culture DNA Midi kit (Qiagen), with final elution of DNA in Tris buffer. Strains included in the analysis were: *Bacillus anthracis* Sterne 24F2 (3ng/uL), *Bacillus thuringiensis* kurstaki (36 ng/uL), *Burkholderia cepacia* ATCC25416 (484 ng/uL), *Escherichia coli* PA03M55684 (60 ng/uL), *Enterococcus faecalis* 159905163 (14 ng/uL), *Staphylococcus aureus* ATCC9144 (14 ng/uL), *Yersinia pseudotuberculosis* ATCC 13979 (324 ng/uL) (Table 1). All DNA was of high quality (Nanodrop 2000 –Thermo Scientific & Bioanalyzer—Agilent) and concentrations were determined using Qubit (ThermoFisher). Pure DNA was then used to create pools of mixed bacterial communities for sequencing as described in Table 1. Group 2 also sequenced each of the pure gDNA samples from the individual isolates, with this data used to explore the rate at

Table 1. Mixed unknown samples used in the comparative community analysis. (Genomic equivalents in millions).

Pool A	Pool B	Pool C	Pool D
<i>B. cepacia</i> (224M)	<i>S. aureus</i> (46M)	<i>Y. pseudotuberculosis</i> (30M)	<i>Y. pseudotuberculosis</i> (61M)
<i>Y. pseudotuberculosis</i> (150M)	<i>E.coli</i> (36M)	<i>B. cepacia</i> (26M)	<i>E. coli</i> (12M)
<i>B. thuringiensis</i> (67M)	<i>B. cepacia</i> (26M)	<i>E.coli</i> (12M)	<i>B. thuringiensis</i> (2.8M)
<i>E. faecalis</i> (32M)	<i>B. anthracis</i> (16M)	<i>B. thuringiensis</i> (5.6M)	<i>B. anthracis</i> (2.5M)
<i>V. parahaemolyticus</i> (19M)		<i>S. aureus</i> (4.6M)	<i>B. cepacia</i> (1.3M)
		<i>E. faecalis</i> (3.8M)	
		<i>V. parahaemolyticus</i> (1.8M)	
		<i>B. anthracis</i> (0.8M)	

<https://doi.org/10.1371/journal.pone.0295028.t001>

which data was generated over time. Groups 2 and 3 were blinded to the composition of samples.

Samples were processed as described in [S1 Fig](#). In brief, each group applied their best practice library preparation protocols and sequence analysis methods to the prepared samples. Sequencing library preparation kits used by each of the groups included the Rapid barcoding (SQK-RBK004; RSE_9046_v1_revW_14Aug2019), and Rapid PCR barcoding kits (SQK-RPB004; RPB_9059_V1_REVA_08MAR2018)(Oxford Nanopore Technologies, Oxford UK). Sequencing was performed using ONT MinION flowcells (R9.4.1) and Mk1B sequencing devices. Flowcells and reagents used by Groups 2 and 3 were transported to the location of the field exercise via commercial aircraft. As such, while an attempt was made to maintain equipment and reagents at appropriate temperatures, they may have been exposed to prolonged thawing/warming and changes in pressure.

Flowcells were assessed for and passed minimum QC metrics at the time of sequencing. Library preparation and sequencing were carried out as per protocols provided by ONT. As this exercise was designed to simulate sequencing in the field, minimal protocols were conducted: no additional DNA quantification was performed by any of the groups prior to library preparation, rather each sample was added using the maximum volume as indicated by the ONT protocol, and AMPure bead cleanup was omitted following barcoding.

Both Groups 1 and 3 made use of Windows-based software products produced and marketed by ONT on systems described in [Table 2](#). Group 2 performed their analysis on an Intel Server, with customized software for data analysis, which included use of albacore (v 2.3.1) (ONT, Oxford UK) and Kraken (1.0) with a reference database consisting of bacterial, viral, archaeal, fungal and protist species, obtained from NCBI's RefSeq collection (2018-April-20) ([Table 2](#)) [10]. An abundance threshold for detection and reporting of 1% is used by default by the What's in my pot (WIMP) analysis pipeline, as such although not part of standard protocol and in order to maintain consistency between methods, Group 2 also applied an abundance threshold cutoff of 1% to the taxonomic analysis.

Results

Each of the groups were able to successfully sequence all of the mixed microbial community samples included as part of this analysis. Group 3 had challenges with getting runs started, due

Table 2. Comparison of computational resources used by each of the participating partners for data capture and analysis.

	Recommended by ONT	Group 1	Group 2	Group 3	MinIT (for comparison)
Platform	Not specified	Intel NUC	LeNovo ThinkServer	Dell Latitude 5580	MinIT
Operating system	Windows/Ubuntu/Sierra	Windows	Ubuntu	Windows	Linux
Memory	16GB RAM	32GB RAM—max	504GB RAM	16G RAM	8GB RAM/GPU accelerators
CPU	i7 or Xeon with 4 + cores	i7, 3.5GHz—max	Intel Xeon E5- 2.20 GHz	Intel R Core i7-7820HQ	Unclear
Storage	1TB SSD	2TB max	3TB (mirrored)	1TB	512GB
USB3 Ports	Yes	Yes	Yes	Yes	Yes
Internet connection	Required	Required for WIMP only	Not Required (ping free)	Required for MinKnow and WIMP (need to ping)	Not Required
Basecalling Software	MinKNOW/Albacore	MinKNOW (2.2)	Albacore (2.3.0)	MinKNOW (2.2)	MinKNOW
Bacterial identification software	WIMP	WIMP (3.2.0)	Kraken (1.0)—custom database; signature sequence analysis	WIMP (3.2.0)	Optional

<https://doi.org/10.1371/journal.pone.0295028.t002>

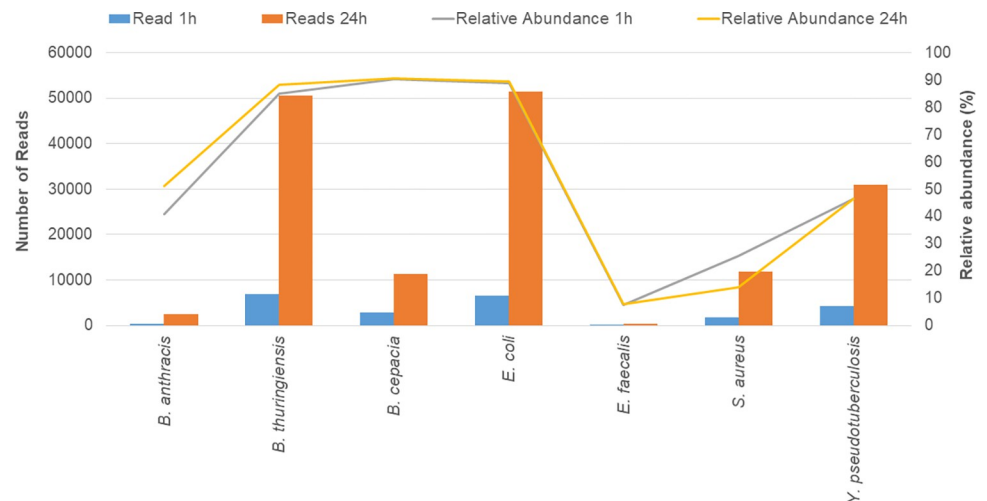


Fig 1. Comparison of identification of species from the pure DNA extracts 1 hour and 24 hours following the start of sequencing, performed by Group 2.

<https://doi.org/10.1371/journal.pone.0295028.g001>

to unstable internet connectivity, which is a requirement of the off-the-shelf MinKNOW software. This group also ran into challenges with data storage and slow basecalling, leading to a reduced amount of data being generated. It is unclear whether these challenges arose due to their computational setup, in which they were just meeting the minimum requirements suggested by ONT, or due to unstable internet connectivity and a lack of access to ping-free software.

Analysis of results obtained 1 and 24 hours following the start of sequencing of the single organism samples (Fig 1), demonstrated that DNA sequence could be assigned to species one hour following the beginning of sequencing. Importantly, the abundance of accurately classified organisms detected in the sample did not substantially change between time points. As sequences are generated immediately upon sample addition to the sequencer, the limitation in time to detection is in the transfer of data to the analysis cloud for samples analyzed using the Epi2Me WIMP workflow (Oxford Nanopore Technologies), and in the time taken to both basecall and taxonomically assign reads, rather than the sequencing itself. This suggests that advancements in computational approaches will reduce the time to detection of organisms of interest.

Results of the sequence analysis of the mixed bacterial pools analyzed by each of the groups after 24 hours of sequencing are included in Fig 2. Group 1 generated a large number of sequences throughout the run, and despite use of the Rapid sequencing protocol, had a similar number of reads detected to that generated by Group 2 when using the Rapid PCR protocol. It is unclear as to the reason for this observation as typically PCR-reliant protocols result in increased sequencing yield, it is possible that the transportation of reagents to the location of the exercise by each of the groups may have led to inconsistent efficiency (Group 2 required international travel, and Group 1 domestic travel). The number of reads generated and analyzed by Group 3 were much lower than either of the other groups, which was attributed to the computational challenges faced by this group, and potentially the transport of equipment and reagents to the exercise location.

All organisms from each of the mixed samples were successfully identified by both groups 1 and 2, although several fell below the 1% abundance threshold used for the analysis. Interestingly, each group observed that community composition varied from the input, with *B.*

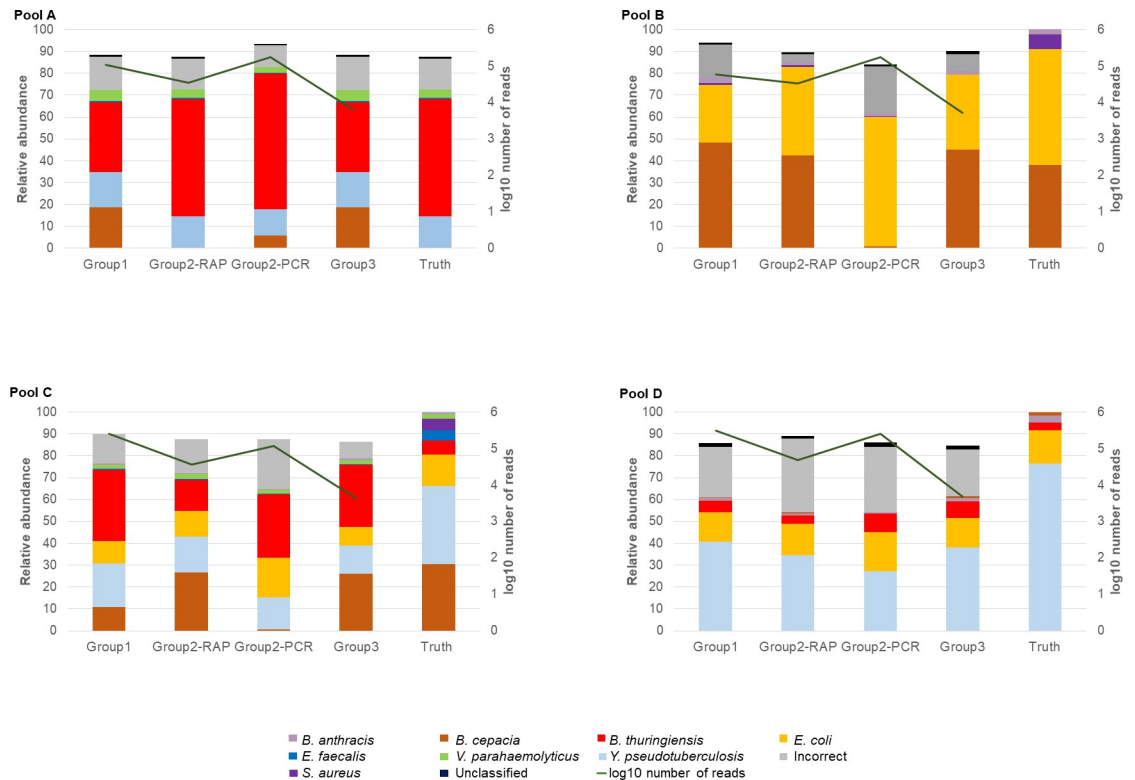


Fig 2. Relative abundance of different species detected in each of the pools by each of the groups following 24 hours of sequencing. Group 2 performed the analysis using both the Rapid (SQK-RBK004—RAP) and Rapid PCR (SQK-RPB004—PCR) barcoding sequencing kits, on separate flowcells. The log₁₀ number of reads detected in each of the samples is described along the right axis. The Rapid PCR sequencing conducted by Group 2 produced so many reads, it was stopped after 8 hours of sequencing.

<https://doi.org/10.1371/journal.pone.0295028.g002>

thuringiensis commonly over represented in samples, and *E. faecalis*, *B. cepacia*, *S. aureus* and *V. parahaemolyticus* commonly under-represented. Notably, *B. cepacia* was detected in a manner more in keeping with the input sample when the Rapid sequencing kit without PCR amplification was used. Incorrect species which were present at >1% relative abundance included both near neighbour species likely arising from both misclassification (*Y. similis* was observed in Pool C by each of the groups, and was observed when Group 2 used both the Rapid and Rapid PCR kits, likely due to its homology with *Y. pseudotuberculosis*) and cross-contamination during the preparation of the sample or sequencing library (*B. thuringiensis* was detected in all pooled samples, including those for which none of this organism was added, and at levels disproportionate to the input amount). Reads which were correctly assigned to taxonomic levels above species, were not included in reporting.

Discussion

The results of this analysis have demonstrated that nanopore-based sequencing has utility in the rapid detection and analysis of both pure and mixed samples in a field deployment scenario, when supported by scientific expertise and appropriate on-site analytical capabilities. This includes, ensuring strong internet connectivity and having available sufficient computing power for identification of sequence reads. While each group approached the analysis of samples and data differently, similar results were obtained, suggesting that consistent results,

especially in the realm of pathogen detection, are possible across methodologies. Importantly, analytical methods used to evaluate microbial composition did not perform substantially differently in terms of output between groups, with user-friendly cloud-based detection methods (WIMP) performing equally to customized pipelines.

An important caveat to this was the experience reported by Group 3, who, while using a computational setup which met the minimum requirements set out by ONT, had trouble both getting runs to start, and produced less sequencing data than did the other groups. Importantly, despite the close physical proximity to other groups, there may have been inconsistent access to wireless internet which hindered this group's ability to perform sequencing. Thus, *a priori* knowledge of the availability of stable internet remains an important consideration for investigators conducting field-based analyses, both at the sequencing stage, and when performing cloud-based analysis. Other tools (MinIT, Mk1C) are capable of sequencing independent of a wireless signal, and are worth investigating for use in scenarios in which internet stability cannot be guaranteed. Alternatively, use of ping free software and locally available computational hardware exceeding the minimum specifications set out by ONT for intensive analyses, also improve reliability and performance. Use of local analysis pipelines also provides users with analytical flexibility, in the event that available protocols do not meet user needs.

Importantly, despite similarities in results between the groups, bacterial quantification in mixed samples by each group was disparate from true sample composition, with no methodology exactly replicating the proportion of reads that were in the initial sample. Such findings are consistent with microbiome sequencing performed by other methods [11], and are speculated to primarily arise from extraction, amplification and sequencing bias produced by the experimental process [12,13]. Such biases are important to consider in designing appropriate universal detection approaches, as pathogenic organisms may be missed when inappropriate analyses are used. This demonstrates that while metagenomic sequencing approaches may be universal in design, in actuality there are limitations which must be considered prior to experimental analysis when hoping to utilize a one-size-fits-all approach. All kits are not created equal and an understanding of the limitations and bias of each is vital to inform operational utility. This is particularly relevant in the event that such methods might be used by field responders who may not have expertise or understanding in the fundamental use of sequencing tools for detection and identification purposes. In addition, including techniques such as genomic or targeted amplification or multiple extractions to target different classes of bacteria may also be beneficial to increase sensitivity [14–17], but have the disadvantage of increasing turn-around-time and being less easily deployable in the field. Perhaps unsurprisingly, sequencing methods which involve less experimental manipulation (rapid sequencing kit–SQK-RBK004) appear to generate results which more closely approximate the input community structure. Thus, in cases in which abundant DNA is available, availability of a sequencing method which does not require additional manipulation or PCR is useful for detecting species which are less amenable to amplification.

While this study was conducted on a small, well-defined dataset, it has highlighted the utility of MinION sequencing in the field, for presumptive identification of bacterial species in mixed samples. Further, applying this as a rapid triage tool to understand microbial composition within a sample can direct the use of targeted, confirmatory testing while in a field scenario. While our results suggest that experimental outcomes are robust to different methodologies, it is important to note that in each case, operation of equipment, and interpretation of data by experienced laboratory personnel remains imperative. Importantly, this work adds to the body of literature supporting the use of ONT's MinION sequencing device for field-based projects, particularly in the context of rapid pathogen identification.

Conclusion

The data presented in this manuscript illustrates the utility of nanopore-based sequencing for in-field detection of bacterial pathogens in mixed samples. The results of this work emphasize that despite all species present in the input being detected in the output sequencing data, alterations in the observed structure of the microbial community were ubiquitous and could confound interpretation by untrained users. Not surprisingly, sequencing methods which did not rely on PCR amplification for DNA analysis showed less sequencing bias than did methods which employed amplification approaches. Such observations must be balanced with the need to maximize sequencing output.

Supporting information

S1 Fig. Workflow depicting analytic strategy used for this analysis.
(PDF)

Author Contributions

Formal analysis: Andrea D. Tyler, Jane McAllister, Helen Stapleton.

Funding acquisition: Cindi R. Corbett.

Methodology: Jane McAllister, Helen Stapleton, Penny Gauci, Kym Antonation.

Project administration: Kym Antonation.

Resources: Jane McAllister, Helen Stapleton, David Thirkettle-Watts, Cindi R. Corbett.

Software: Andrea D. Tyler.

Supervision: Cindi R. Corbett.

Writing – original draft: Andrea D. Tyler.

Writing – review & editing: Jane McAllister, Helen Stapleton, Penny Gauci, Kym Antonation, Cindi R. Corbett.

References

1. Tyler AD, Mataseje L, Urfano CJ, Schmidt L, Antonation KS, Mulvey MR, et al. Evaluation of Oxford Nanopore's MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. *Sci Rep* [Internet]. 2018 Dec 19 [cited 2018 Dec 13]; 8(1):10931. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30026559>. <https://doi.org/10.1038/s41598-018-29334-5> PMID: 30026559
2. Parker J, Helmstetter AJ, Devey D, Wilkinson T, Papadopoulos AST. Field-based species identification of closely-related plants using real-time nanopore sequencing. *Sci Rep* [Internet]. 2017 Dec 21 [cited 2018 Dec 13]; 7(1):8345. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28827531>. <https://doi.org/10.1038/s41598-017-08461-5> PMID: 28827531
3. Leidenfrost RM, Pöther DC, Jäckel U, Wünschiers R. Benchmarking the MinION: Evaluating long reads for microbial profiling. *Sci Rep* [Internet]. 2020 Dec 1 [cited 2021 Feb 1]; 10(1):1–10. Available from: <https://doi.org/10.1038/s41598-020-61989-x>.
4. Walter MC, Zwirgmaier K, Vette P, Holowachuk SA, Stoecker K, Genzel GH, et al. MinION as part of a biomedical rapidly deployable laboratory. *J Biotechnol* [Internet]. 2017 May 20 [cited 2017 Aug 18]; 250:16–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27939320>. <https://doi.org/10.1016/j.jbiotec.2016.12.006> PMID: 27939320
5. Bates M, Polepole P, Kapata N, Loose M, O'Grady J. Application of highly portable MinION nanopore sequencing technology for the monitoring of nosocomial tuberculosis infection. *Int J Mycobacteriology* [Internet]. 2016 Dec [cited 2017 Jul 14]; 5:S24. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28043580>. <https://doi.org/10.1016/j.ijmyco.2016.10.035> PMID: 28043580
6. 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*

- [Internet]. 2015 Dec 29 [cited 2017 Sep 25]; 7(1):99. Available from: <http://genomemedicine.biomedcentral.com/articles/10.1186/s13073-015-0220-9>. <https://doi.org/10.1186/s13073-015-0220-9> PMID: 26416663
7. Goordial J, Altshuler I, Hindson K, Chan-Yam K, Marcollef E, Whyte LG. In Situ Field Sequencing and Life Detection in Remote (79°26'N) Canadian High Arctic Permafrost Ice Wedge Microbial Communities. *Front Microbiol* [Internet]. 2017 Dec 20 [cited 2019 May 9]; 8. Available from: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.02594/full>.
 8. Pontefract A, Hachey J, Zuber MT, Ruvkun G, Carr CE. Sequencing nothing: Exploring failure modes of nanopore sensing and implications for life detection. *Life Sci Sp Res* [Internet]. 2018 Aug [cited 2019 May 9]; 18:80–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30100151>. <https://doi.org/10.1016/j.lssr.2018.05.004> PMID: 30100151
 9. Krishnakumar R, Sinha A, Bird SW, Jayamohan H, Edwards HS, Schoeniger JS, et al. Systematic and stochastic influences on the performance of the MinION nanopore sequencer across a range of nucleotide bias. *Sci Rep* [Internet]. 2018 Dec 16 [cited 2019 May 9]; 8(1):3159. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29453452>. <https://doi.org/10.1038/s41598-018-21484-w> PMID: 29453452
 10. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* [Internet]. 2014 [cited 2017 Jul 21]; 15(3):R46. Available from: <http://genomebiology.biomedcentral.com/articles/10.1186/gb-2014-15-3-r46>. <https://doi.org/10.1186/gb-2014-15-3-r46> PMID: 24580807
 11. de Muinck EJ, Trosvik P, Gilfillan GD, Hov JR, Sundaram AYM. A novel ultra high-throughput 16S rRNA gene amplicon sequencing library preparation method for the Illumina HiSeq platform. *Microbiome* [Internet]. 2017 Dec 6 [cited 2019 May 10]; 5(1):68. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28683838>. <https://doi.org/10.1186/s40168-017-0279-1> PMID: 28683838
 12. Kim D, Hofstaedter CE, Zhao C, Mattei L, Tanes C, Clarke E, et al. Optimizing methods and dodging pitfalls in microbiome research [Internet]. Vol. 5, *Microbiome*. BioMed Central Ltd.; 2017 [cited 2021 Mar 15]. p. 1–14. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-017-0267-5>.
 13. Browne PD, Nielsen TK, Kot W, Aggerholm A, Gilbert MTP, Puetz L, et al. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. *Gigascience* [Internet]. 2020 Feb 1 [cited 2021 Mar 15]; 9(2):1–14. Available from: <http://orcid.org/0000-0001-8300-7758andpdr@plen.ku.dkhttp://orcid.org/0000-0003-0027-1524>. <https://doi.org/10.1093/gigascience/giaa008> PMID: 32052832
 14. Kim D, Hofstaedter CE, Zhao C, Mattei L, Tanes C, Clarke E, et al. Optimizing methods and dodging pitfalls in microbiome research [Internet]. Vol. 5, *Microbiome*. BioMed Central Ltd.; 2017 [cited 2021 Apr 9]. p. 1–14. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-017-0267-5>.
 15. Wesolowska-Andersen A, Bahl MI, Carvalho V, Kristiansen K, Sicheritz-Pontén T, Gupta R, et al. Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome* [Internet]. 2014 Jun 5 [cited 2021 Apr 9]; 2(1):19. Available from: <http://www.microbiomejournal.com/content/2/1/19>. <https://doi.org/10.1186/2049-2618-2-19> PMID: 24949196
 16. Andreani J, Million M, Baudoin JP, Ominami Y, Khalil JYB, Frémond C, et al. Klenkia terrae resistant to DNA extraction in germ-free mice stools illustrates the extraction pitfall faced by metagenomics. *Sci Rep* [Internet]. 2020 Dec 1 [cited 2021 Apr 9]; 10(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/32576848/>. <https://doi.org/10.1038/s41598-020-66627-0> PMID: 32576848
 17. Yang Y, Walls SD, Gross SM, Schroth GP, Jarman RG, Hang J. Targeted sequencing of respiratory viruses in clinical specimens for pathogen identification and genome-wide analysis. In: *Methods in Molecular Biology* [Internet]. Humana Press Inc.; 2018 [cited 2021 Apr 9]. p. 125–40. Available from: <https://pubmed.ncbi.nlm.nih.gov/30128994/>. https://doi.org/10.1007/978-1-4939-8682-8_10 PMID: 30128994