Metagenomic Next-Generation Sequencing of Rectal Swabs for the Surveillance of Antimicrobial Resistant Organisms on the Illumina Miseq and Oxford MinION Platforms

Rebecca Yee¹, Florian P. Breitwieser², Stephanie Hao³, Belita N.A. Opene¹, Rachael E. Workman³, Pranita D. Tamma⁴, Jennifer Dien-Bard⁵, Winston Timp³, and Patricia J. Simner¹

1 Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD

2 Center for Computational Biology, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, MD

3 Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD

4 Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD

5 Department of Pathology and Laboratory Medicine, Children's Hospital of Los Angeles and Keck School of Medicine at the University of Southern California, Los Angeles, California, USA

Running title: mNGS of Rectal Swabs for Antimicrobial Resistant Organisms

Key words: metagenomics, next-generation sequencing, antimicrobial resistance, microbiome, resistome, surveillance

Corresponding Author:

Patricia (Trish) J. Simner, PhD, D(ABMM)

Associate Professor of Pathology

Director of Bacteriology and Parasitology, Division of Medical Microbiology

Johns Hopkins Medical Institutions

25 Department of Pathology
26 Division of Medical Microbiology
27 Meyer B1-193, 600 N. Wolfe Street
28 Baltimore, MD 21287-7093
29 Phone: 410-955-5077 / Fax: 410-614-8087
30 psimmer1@jhmi.edu

Abstract

Purpose: Antimicrobial resistance (AMR) is a public health threat where efficient surveillance is needed to prevent outbreaks. Existing methods for detection of gastrointestinal colonization of multidrug-resistant organisms (MDRO) are limited to specific organisms or resistance mechanisms. Metagenomic next-generation sequencing (mNGS) is a more rapid and agnostic diagnostic approach for microbiome and resistome investigations. We determined if mNGS can detect MDRO from rectal swabs in concordance with standard microbiology results.

Methods: We performed and compared mNGS performance on short-read Illumina MiSeq (N=10) and long-read Nanopore MinION (N=4) platforms directly from peri-rectal swabs to detect vancomycin-resistant enterococci (VRE) and carbapenem-resistant Gram-negative organisms (CRO).

Results: We detected *E. faecium* (N=8) and *E. faecalis* (N=2) with associated *van* genes (9/10) in concordance with VRE culture-based results. We studied the microbiome and identified CRO organisms, *P. aeruginosa* (N=1), *E. cloacae* (N=1), and KPC-producing *K. pneumoniae* (N=1). Nanopore real-time detection detected the *bla_{KPC}* gene in 2.5 minutes and provided genetic context (*bla_{KPC}* harbored on pKPC_Kp46 IncF plasmid). Illumina sequencing provided accurate allelic variant determination (i.e., *bla_{KPC-2}*) and strain typing of the *K. pneumoniae* (ST-15).

Conclusions: We demonstrated an agnostic approach for surveillance of MDRO, examining advantages of both short and long-read mNGS methods for AMR detection.

INTRODUCTION

Early detection of colonization by multidrug-resistant organism (MDRO) such as vancomycin-resistant enterococci (VRE) and carbapenem-resistant organisms (CRO) can lead to early implementation of infection prevention practices, antimicrobial optimization, and prevention of invasive infections [1]. Current methods (selective culture techniques or PCR) are targeted towards a specific MDRO [2]. To identify novel mechanisms of resistance or emerging pathogens, a broader approach to detect and characterize MDRO is required.

Metagenomic next-generation sequencing (mNGS) of specimens using next-generation sequencing (NGS) platforms is an agnostic approach. mNGS amplifies any DNA in the sample. Thus, this approach can query the entire microbiome of the sample, and provide valuable information about the resistome (all known antimicrobial resistance genes), and plasmids [3].

In this proof-of-concept study, we determined if mNGS using short-read Illumina sequencing and long-read Oxford Nanopore sequencing can be applied to rectal swabs for detection of VRE and CROs identified by standard microbiology results.

Materials and Methods

Bacterial Isolates and Characterization. Remnant rectal surveillance swabs from 10 ICU patients hospitalized at Johns Hopkins Hospital were evaluated for VRE on VRE Select chromogenic agar and CRO by the Direct MacConkey method [2]. Carbapenemase production was detected by the Carba NP assay and Check-MDR CT103XL from cultured isolates and Check-Direct CPE screen assay for the BD MAX instrument from rectal swabs (Check-Points; Becton Dickinson) [4]. Bacteria were identified by matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (Bruker Daltonic). Positive controls (10^2 , 10^4 , or 10^6 CFU/mL) swabs seeded with *Klebsiella pneumoniae* ATCC BAA-1705 (*bla*_{KPC}), *Enterococcus faecalis*

ATCC 51299 (*vanB*) and *Enterococcus faecium* ATCC 700228 (*vanA*) and negative control swabs were also sequenced. ESwabs were de-identified, frozen at -70°C, and DNA was extracted from the broth (500 µl) using the Zymo ZR Fungal/Bacterial Miniprep Kit.

Illumina Sequencing. Library preparation was performed using Illumina NexteraXT™ DNA Sample Prep Kit per manufacturer’s protocol followed by AMPure XP (Beckman Coulter) purification. Normalized samples were pooled (n=4) and sequenced on a MiSeq v3 2x75 flowcell. Reads were assembled using metaSPAdes [5].

Oxford Nanopore Sequencing. Library preparation was performed using the low-input genomic DNA sequencing kit protocol for SQK-MAP006 per manufacturer’s instructions. Libraries were loaded onto a R7.3 flowcell, sequenced using MinKNOW, and basecalled using Metrichor. Only Nanopore 2d high quality reads were used.

Bioinformatics. Taxonomic analysis was performed with Kraken and plotted using Krona [6]. Resistance genes were queried using BLAST and against ResFinder using Abricate [5]. Alignment to *K. pneumoniae* strain KPNIH49 (RefSeq GCF_002903025.1) was done using bowtie2 and visualized using Pavian [7]. Data analysis and heatmap visualization was done using R statistical environment [8]. Our sequencing data, scrubbed of human reads, has been deposited at SRA (accession pending).

Results

Characterization of Microbiome from Rectal Swabs. The results from both platforms were concordant with standard microbiology culture methods (Table 1). We detected VRE (*E. faecium* and/or *E. faecalis*), *vanA* and/or *vanB* in all samples, and predominant CRO organisms (*E. cloacae*, *P. aeruginosa* and *K. pneumoniae*) in 3 samples (Figure 1A-B). Both platforms detected the same or similar species as the top three dominant organisms (Table 1).

Metagenomic data and frequency of classified reads were similar between both platforms (Figure 1, see web-only Supplementary Table S1). Similar percentages of human reads (range 0.01% and 10%), with Swab #8 as an outlier (44%) were generated. Our negative control had greater than 60% human reads on both platforms.

AMR Gene Detection. MiSeq detected *vanA* in all samples except Swab #4, which showed a lower quantity of growth by culture suggesting that initial sampling may be low. Swabs #2 and #10 had *vanB* detected and Swab #7 had *bla*_{KPC} detected, corresponding to culture and PCR results from the rectal swab. The *bla*_{KPC} gene was detected by both platforms but non-carbapenemase mediated mechanisms of carbapenem resistance were not confirmed by mNGS. Our bioinformatics analyses also revealed AMR genes for other antimicrobial classes (Figure 1). All controls sequenced generated data concordant with the seeded organisms and resistance genes.

Case-study: Correlation of Phenotypic Culture Results to Molecular mNGS Analysis. In Swab #7, we identified vancomycin-resistant *E. faecium* (Figure 1C) and a lactose fermenter with a 12 mm zone of inhibition around an ertapenem disk (Figure 1D), further revealed to be a KPC-producing *K. pneumoniae* by PCR from cultured isolate (Figure 1E). From both platforms, we detected *K. pneumoniae* as the dominant and *E. faecium* as the second organism (Figure 1F). MiSeq further determined the allele as *bla*_{KPC-2}, identifying mutations to the single nucleotide level for allelic variant identification and strain typing of ST-15 for the predominant *K. pneumoniae* directly from specimen (Figure 1G-H). For other AMR genes, the accuracy and percent coverage for allelic variants detected by MinION (96.9%) was lower than MiSeq (99.7%) (Figure 1G). Meanwhile, MinION's ability for rapid real-time analysis of AMR genes detected *bla*_{KPC} in 2.3 minutes. Additionally, MinION sequencing provided the ability to

associate the *bla*_{KPC} gene to the IncFII-type plasmid pKPC_Kp46 [9]. Heatmap of the full resistome was also generated (Figure 1I).

Discussion

The shortcomings of targeted PCRs and the requirement for a priori knowledge of the resistance markers is more evident as additional MDROs are identified, highlighting the potential benefits of mNGS as an alternative approach. Despite its high accuracy (99%) in strain typing and AMR allelic variant determination, Illumina sequencing may not be actionable for clinical care usages due to long run times of 24-48 hours. Nanopore sequencing offers advantages of real-time base-calling, detection of resistance genes in as little as 2 minutes and determination of the genetic context of the AMR genes detected. A real-time WGS approach detected all AMR genes within 14 minutes and can shorten time to effective therapy for carbapenem-resistant *K. pneumoniae* infections by 20 h compared to standard approaches [5]. In 6 hours, full annotation of plasmid-based resistance genes was achieved in extended-spectrum β -lactamase-producing *E. coli* and *K. pneumoniae* isolates [10].

Our study here is one of the first to compare platforms for mNGS on rectal swabs where accurate MDRO and resistance genes were identified compared to culture-based methods. Mu *et al* identified a KPC-producing *K. pneumoniae* isolate using MiSeq but only tested one rectal swab [11]. The similar performance of both platforms seen in our study has been observed in others; comparable phylogenetic trees for *N. gonorrhoeae* [12] and detection of Dengue and Chikungunya viruses from plasma and serum samples [13] have been shown. Similarly, Nanopore sequencing distinguished allelic variants poorly by flagging multiple alleles (i.e., *bla*_{KPC-2}, *bla*_{KPC-3}, etc.) while Illumina sequencing detected single variants (i.e., *bla*_{KPC-2}) and performed strain typing directly from specimens [14].

Limitations include the lack of broader species and resistance mechanisms tested. Newer sequencing methods (rapid extraction and library kits, Flongle) and flow cells (MinION R.9 and R.10) from Oxford Nanopore may improve accuracy significantly [15].

In conclusion, mNGS analysis may be a promising approach for detection of MDRO from rectal swabs. mNGS allows the study of the entire microbiome, providing important clinical information such as the resistome, allelic variants, and strain typing to guide infection control and patient management. Future studies evaluating newer technologies and automated processes are necessary to increase the efficiency and advance mNGS methods for patient care.

References

1. CDC (2019) Antibiotic Resistance Threats In The United States, 2019. CDC,
2. Simner PJ, Martin I, Opene B, Tamma PD, Carroll KC, Milstone AM (2016) Evaluation of Multiple Methods for Detection of Gastrointestinal Colonization of Carbapenem-Resistant Organisms from Rectal Swabs. *Journal of clinical microbiology* 54 (6):1664-1667. doi:10.1128/jcm.00548-16
3. Li X, Arias CA, Aitken SL, Galloway Pena J, Panesso D, Chang M, Diaz L, Rios R, Numan Y, Ghaoui S, DebRoy S, Bhatti MM, Simmons DE, Raad I, Hachem R, Folan SA, Sahasrabhojane P, Kalia A, Shelburne SA (2018) Clonal Emergence of Invasive Multidrug-Resistant *Staphylococcus epidermidis* Deconvoluted via a Combination of Whole-Genome Sequencing and Microbiome Analyses. *Clin Infect Dis*. doi:10.1093/cid/ciy089
4. Clinical and Laboratory Standards Institute. (2019) Performance Standards for Antimicrobial Susceptibility Testing; Twenty-Ninth Informational Supplement., vol M100-S28.
5. Tamma PD, Fan Y, Bergman Y, Perteza G, Kazmi AQ, Lewis S, Carroll KC, Schatz MC, Timp W, Simner PJ (2019) Applying Rapid Whole-Genome Sequencing To Predict Phenotypic Antimicrobial Susceptibility Testing Results among Carbapenem-Resistant *Klebsiella pneumoniae* Clinical Isolates. *Antimicrobial agents and chemotherapy* 63 (1). doi:10.1128/aac.01923-18
6. Wood DE, Salzberg SL (2014) Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome biology* 15 (3):R46. doi:10.1186/gb-2014-15-3-r46
7. Breitwieser FP, Salzberg SL (2019) Pavian: Interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics*. doi:10.1093/bioinformatics/btz715
8. R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
9. Kim JO, Song SA, Yoon EJ, Shin JH, Lee H, Jeong SH, Lee K (2017) Outbreak of KPC-2-producing Enterobacteriaceae caused by clonal dissemination of *Klebsiella pneumoniae* ST307 carrying an IncX3-type plasmid harboring a truncated Tn4401a. *Diagnostic microbiology and infectious disease* 87 (4):343-348. doi:10.1016/j.diagmicrobio.2016.12.012
10. Lemon JK, Khil PP, Frank KM, Dekker JP (2017) Rapid Nanopore Sequencing of Plasmids and Resistance Gene Detection in Clinical Isolates. *Journal of clinical microbiology* 55 (12):3530-3543. doi:10.1128/jcm.01069-17
11. Mu A, Kwong JC, Isles NS, Goncalves da Silva A, Schultz MB, Ballard SA, Lane CR, Carter GP, Williamson DA, Seemann T, Stinear TP, Howden BP (2019) Reconstruction of the Genomes of Drug-Resistant Pathogens for Outbreak Investigation through Metagenomic Sequencing. *mSphere* 4 (1). doi:10.1128/mSphere.00529-18
12. Golparian D, Dona V, Sanchez-Buso L, Foerster S, Harris S, Endimiani A, Low N, Unemo M (2018) Antimicrobial resistance prediction and phylogenetic analysis of *Neisseria gonorrhoeae* isolates using the Oxford Nanopore MinION sequencer. *Scientific reports* 8 (1):17596. doi:10.1038/s41598-018-35750-4
13. Kafetzopoulou LE, Efthymiadis K, Lewandowski K, Crook A, Carter D, Osborne J, Aarons E, Hewson R, Hiscox JA, Carroll MW, Vipond R, Pullan ST (2018) Assessment of metagenomic Nanopore and Illumina sequencing for recovering whole genome sequences of chikungunya and dengue viruses directly from clinical samples. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 23 (50). doi:10.2807/1560-7917.es.2018.23.50.1800228
14. Schmidt K, Mwaigwisya S, Crossman LC, Doumith M, Munroe D, Pires C, Khan AM, Woodford N, Saunders NJ, Wain J, O'Grady J, Livermore DM (2017) Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based

214 metagenomic sequencing. The Journal of antimicrobial chemotherapy 72 (1):104-114.
 215 doi:10.1093/jac/dkw397
 216 15. Nicholls SM, Quick JC, Tang S, Loman NJ (2019) Ultra-deep, long-read nanopore
 217 sequencing of mock microbial community standards. GigaScience 8 (5).
 218 doi:10.1093/gigascience/giz043
 219

220 **Table 1:** Summary of Culture Based Results for the Detection of Antimicrobial Resistant Organisms from Rectal Swabs Compared to
221 mNGS on the MiSeq and MinION Platforms

Study #	Vancomycin-resistant enterococci (VRE) Culture Results		Carbapenem-Resistant Organism Culture Results		CheckDirect CPE Results	Miseq Data		Nanopore Data	
	VRE culture results	Organism based on chromogen	Gram-negative growth on MacConkey Agar with ertapenem disks ^a	Carbapenem resistant organism (CRO) culture results	Detection of carbapenemase genes directly from rectal swabs	Dominant Bacterial organisms	Detection of <i>vanA</i> , <i>vanB</i> or <i>bla</i> _{KPC}	Dominant Bacterial Organisms ^b	Detection of <i>vanA</i> , <i>vanB</i> or <i>bla</i> _{KPC} ^b
1	Positive for VRE	<i>E. faecium</i>	4 + Lactose fermenter; >27mm zone diameter	Negative	Negative	<i>Klebsiella pneumoniae</i> , <i>Clostridioides difficile</i> , <i>Enterococcus faecium</i>	<i>vanA</i> positive	N/A	N/A
2	Positive for VRE	<i>Enterococcus faecalis</i>	4+ Mixed lactose fermenters; >27mm zone diameter	Negative	Negative	<i>Enterobacter cloacae</i> , <i>Parabacteroides distasonis</i> , <i>Enterobacter asburiae</i>	<i>vanA</i> and <i>vanB</i> positive	N/A	N/A
3	Positive for VRE	<i>E. faecium</i>	4+ Non-lactose fermenter; no zone around ertapenem disk	Positive: Meropenem resistant <i>Pseudomonas aeruginosa</i>	Negative	<i>Pseudomonas aeruginosa</i> , <i>Enterococcus faecalis</i> , <i>Enterococcus faecium</i>	<i>vanA</i> positive	<i>Pseudomonas aeruginosa</i> , <i>Enterococcus faecalis</i> , <i>Enterococcus faecium</i>	<i>vanA</i> positive
4	Positive for VRE	<i>E. faecium</i>	4+ Non-lactose fermenter; 21 mm zone around ertapenem disk	Positive: Ertapenem resistant <i>Enterobacter cloacae</i>	Negative	<i>Enterobacter cloacae</i> , <i>Enterobacter asburiae</i> , <i>Cronobacter sakazakii</i>	no detection	N/A	N/A

5	Positive for VRE	<i>E. faecium</i>	No Growth	Negative	Negative	<i>Enterococcus faecium</i> , <i>Corynebacterium jeikeium</i> , <i>Parabacteroides</i> spp.	<i>vanA</i> positive	<i>Enterococcus faecium</i> , <i>Corynebacterium jeikeium</i> , <i>Parabacteroides</i> spp.	<i>vanA</i> positive
6	Positive for VRE	<i>E. faecium</i>	1+ Non-lactose fermenter; >40 mm zone around ertapenem disk	Negative	Negative	<i>Achromobacter xylosoxidans</i> , <i>Enterococcus faecium</i> , <i>Pseudomonas aeruginosa</i> ,	<i>vanA</i> positive	<i>Achromobacter xylosoxidans</i> , <i>Enterococcus faecium</i> , <i>Bacillus cereus</i>	<i>vanA</i> positive
7	Positive for VRE	<i>E. faecium</i>	1+ Lactose fermenter; 12 mm zone around ertapenem disk	Positive: KPC-producing <i>K. pneumoniae</i>	<i>bla</i> _{KPC} Ct: 14.1	<i>Klebsiella pneumoniae</i> , <i>Proteus mirabilis</i> , <i>Enterococcus faecium</i>	<i>vanA</i> , <i>bla</i> _{KPC} positive	<i>Klebsiella pneumoniae</i> , <i>Enterococcus faecium</i> , <i>Bacillus cereus</i>	<i>vanA</i> , <i>bla</i> _{KPC} positive
8	Positive for VRE	<i>E. faecium</i>	Mixed Lactose fermenters; 36 mm zone around ertapenem disk	Negative	Negative	<i>Klebsiella pneumoniae</i> , <i>Enterococcus faecalis</i> , <i>Enterococcus faecium</i> ,	<i>vanA</i> positive	N/A	N/A
9	Positive for VRE	<i>E. faecium</i>	No Growth	Negative	Negative	<i>Enterococcus faecium</i> , <i>Enterococcus faecalis</i> , <i>Corynebacterium jeikeium</i> ,	<i>vanA</i> positive	N/A	N/A
10	Positive for VRE	<i>E. faecalis</i>	Non-lactose fermenter ("Pseudo"); 39 mm zone around ertapenem	Negative	Negative	<i>Parabacteroides distasonis</i> , <i>Enterobacter cloacae</i> , <i>Enterobacter hormaechei</i>	<i>vanA</i> , <i>vanB</i> , <i>bla</i> _{KPC} positive ^c	<i>Parabacteroides distasonis</i> , <i>Enterobacter cloacae</i> , <i>Klebsiella pneumoniae</i>	<i>vanB</i> positive

222 ^a 1-4+ indicates the amount of growth on the plates. 1+: growth in first quadrant, 2+: is growth in the second quadrant, 3+ is growth

223 into the third quadrant and 4+ is growth into the 4th quadrant.

224 ^b N/A: Not applicable as swabs 1, 2, 4, 8 and 9 were not run on the Nanopore platform

225 ^c While *bla*_{KPC} was detected in Swab #10, we did not detect *bla*_{KPC} using traditional microbiologic methods. Detection was likely due

226 to errors in de-multiplexing, as Swabs #7 and #10 were ran on the same flowcell. The low number of reads (n=3) for the *bla*_{KPC}

227 gene in Swab #10 suggests that the detection was likely an artifact.

228

Figure 1

Figure 1 (continued)

Fig. 1 Microbiome and Resistome Analyses Performed on Rectal Swabs. Depiction of organisms with the most percentage of reads from all ten swabs tested on (A) Illumina MiSeq and selected swabs tested on (B) Oxford Nanopore MinION. Taxonomic analysis was performed with Kraken using kmer sizes of 24. Positive controls (PC1, PC2, PC3) were spiked with varying proportions (10^2 , 10^4 or 10^6 CFU/mL) of the following organisms *Klebsiella pneumoniae* (*bla*_{KPC} positive; KPN), *Enterococcus faecalis* (*vanB* positive; ENFS) and *Enterococcus faecium* (*vanA* positive; ENFM). Organisms used for the positive control spikes were also sequenced individually. Negative control (NC) was a pool of rectal E-swabs found to be negative for VRE and CRO by culture. Culture-based methods showed rectal swab #7 contained a (C) positive vancomycin-resistance *Enterococcus faecium* on VRE Select chromogenic agar and (D) 1+ growth of a lactose fermenter producing a 12 mm zone of inhibition around an ertapenem disk on MacConkey agar. (E) Carbapenemase gene *bla*_{KPC} was detected with a Ct value of 14.1 using CheckDirect CPE screen assay directly from the rectal swab. (F) Krona plot of metagenomics analyses from sequencing performed on Illumina MiSeq revealed *K. pneumoniae* as the dominant organism and also detection of *E. faecium* as the second organism, both of which were detected by culture-based methods. (G) Coverage comparison of short-read Illumina MiSeq and long-read Oxford Nanopore MinION revealed higher accuracy and coverage on Illumina MiSeq. (H) Accuracy of Illumina reads allowed for straining typing of the KPC-producing *K. pneumoniae* ST-15. (I) Resistome analyses demonstrated the presence of both vancomycin (*vanA* operon) and carbapenem resistance (*bla*_{KPC}) genes based on adaptation of using ResFinder and additional BLAST analyses