

RESEARCH

Open Access



Metagenomic prediction of antimicrobial resistance in critically ill patients with lower respiratory tract infections

Paula Hayakawa Serpa^{1,2†}, Xianding Deng^{3†}, Mazin Abdelghany¹, Emily Crawford^{2,4}, Katherine Malcolm⁵, Saharai Caldera^{1,2}, Monica Fung¹, Aaron McGeever², Katrina L. Kalantar⁶, Amy Lyden², Rajani Ghale^{1,5}, Thomas Deiss^{1,5}, Norma Neff², Steven A. Miller³, Sarah B. Doernberg¹, Charles Y. Chiu³, Joseph L. DeRisi^{2,7}, Carolyn S. Calfee⁴ and Charles R. Langelier^{1,2*}

Abstract

Background: Antimicrobial resistance (AMR) is rising at an alarming rate and complicating the management of infectious diseases including lower respiratory tract infections (LRTI). Metagenomic next-generation sequencing (mNGS) is a recently established method for culture-independent LRTI diagnosis, but its utility for predicting AMR has remained unclear. We aimed to assess the performance of mNGS for AMR prediction in bacterial LRTI and demonstrate proof of concept for epidemiological AMR surveillance and rapid AMR gene detection using Cas9 enrichment and nanopore sequencing.

Methods: We studied 88 patients with acute respiratory failure between 07/2013 and 9/2018, enrolled through a previous observational study of LRTI. Inclusion criteria were age ≥ 18 , need for mechanical ventilation, and respiratory specimen collection within 72 h of intubation. Exclusion criteria were decline of study participation, unclear LRTI status, or no matched RNA and DNA mNGS data from a respiratory specimen. Patients with LRTI were identified by clinical adjudication. mNGS was performed on lower respiratory tract specimens. The primary outcome was mNGS performance for predicting phenotypic antimicrobial susceptibility and was assessed in patients with LRTI from culture-confirmed bacterial pathogens with clinical antimicrobial susceptibility testing ($n = 27$ patients, $n = 32$ pathogens). Secondary outcomes included the association between hospital exposure and AMR gene burden in the respiratory microbiome ($n = 88$ patients), and AMR gene detection using Cas9 targeted enrichment and nanopore sequencing ($n = 10$ patients).

Results: Compared to clinical antimicrobial susceptibility testing, the performance of respiratory mNGS for predicting AMR varied by pathogen, antimicrobial, and nucleic acid type sequenced. For gram-positive bacteria, a combination of RNA + DNA mNGS achieved a sensitivity of 70% (95% confidence interval (CI) 47–87%) and specificity of 95% (CI 85–99%). For gram-negative bacteria, sensitivity was 100% (CI 87–100%) and specificity 64% (CI 48–78%). Patients with hospital-onset LRTI had a greater AMR gene burden in their respiratory microbiome versus those with

[†]Paula Hayakawa Serpa and Xianding Deng contributed equally to this work.

*Correspondence: chaz.langelier@ucsf.edu

² Chan Zuckerberg Biohub, San Francisco, CA, USA

Full list of author information is available at the end of the article



community-onset LRTI ($p = 0.00030$), or those without LRTI ($p = 0.0024$). We found that Cas9 targeted sequencing could enrich for low abundance AMR genes by > 2500-fold and enabled their rapid detection using a nanopore platform.

Conclusions: mNGS has utility for the detection and surveillance of resistant bacterial LRTI pathogens.

Background

Antimicrobial resistance (AMR) presents a clear threat to human health and is responsible for increasing rates of treatment failure in patients with lower respiratory tract infections (LRTI), the leading cause of infectious disease-related mortality [1]. Implementing effective and targeted therapies in patients with LRTI necessitates not only accurate detection of a broad range of pathogens, but also requires assessment of their resistance to antimicrobials. In many cases, assessment of AMR is not possible due to the need to first isolate a bacterial pathogen in culture prior to antimicrobial susceptibility testing (AST), a process that can require several days and have low yield in the setting of prior antibiotic use [2, 3]. In the absence of a definitive microbiologic diagnosis, LRTI treatment is by necessity empiric, which leads to broad-spectrum antibiotic overuse and selects for resistant pathogens [4, 5].

Metagenomic next-generation sequencing (mNGS) holds promise for overcoming the limitations of traditional respiratory diagnostics by affording culture-independent detection of pathogens and simultaneous profiling of host gene expression signatures of infection [6]. In principle, mNGS can also be used to predict pathogen AMR by detecting bacterial resistance genes. While the performance of cultured bacterial isolate whole genome sequencing has been extensively characterized [7], studies assessing the performance of direct respiratory specimen mNGS for predicting AMR have remained more limited [8–12].

This is in part due to the low abundance of pathogen AMR genes in respiratory and other clinical body fluids, which challenges their detection using conventional mNGS methods [12]. Recent work has demonstrated the potential for CRISPR/Cas9 targeted enrichment using FLASH (Finding Low Abundance Sequences by Hybridization) to overcome this challenge by enhancing detection of low abundance AMR genes in clinical samples. Independent validation of FLASH in a clinical cohort, however, has been needed.

Here, we address these gaps by studying a cohort of critically ill patients to assess the potential of both DNA and RNA mNGS to predict LRTI bacterial pathogen

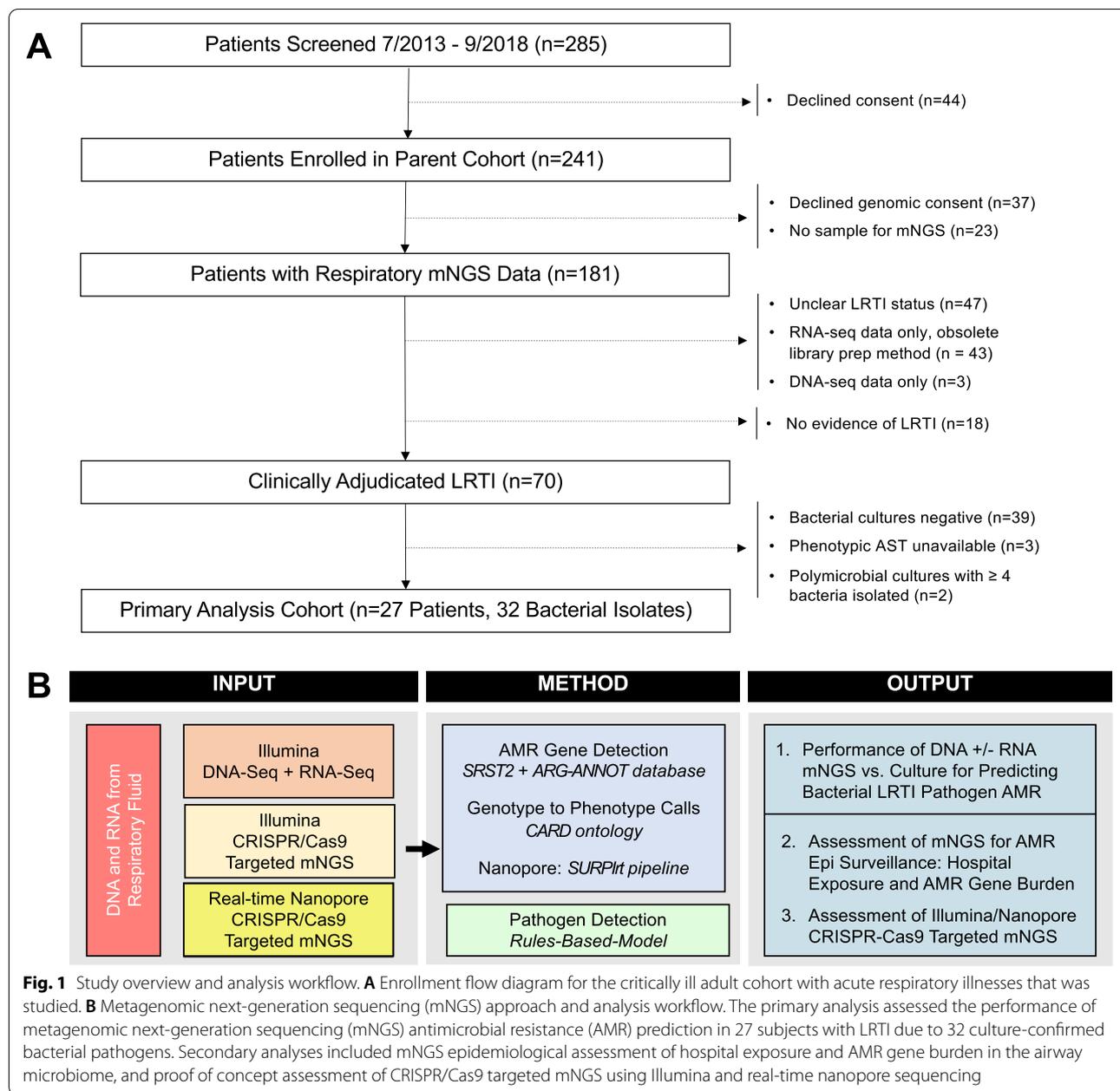
AMR, facilitate epidemiological AMR surveillance, and rapidly detect clinically relevant resistance genes using CRISPR/Cas9 targeted enrichment coupled with real-time nanopore sequencing.

Methods

Study design

We studied 70 mechanically ventilated patients with LRTI and 18 with non-infectious respiratory illnesses (Fig. 1, Additional File 1: Table S1) who were admitted to the University of California San Francisco (UCSF) Medical Center between 07/2013 and 9/2018. Subjects with LRTI were identified by two-physician adjudication using the United States Centers for Disease Control/National Healthcare Safety Network (CDC/NHSN) surveillance case definition [13], a reference list of established respiratory pathogens [6], and retrospective electronic medical record review, blinded to mNGS results. Study inclusion criteria were age ≥ 18 , need for mechanical ventilation, and lower respiratory specimen (tracheal aspirate (TA) or mini-bronchoalveolar lavage (mBAL)) collected within 72 h of intubation. Patients were excluded if they declined study participation, had unclear LRTI status, or did not have matched RNA and DNA mNGS data available from a respiratory specimen (Fig. 1).

Primary analyses were performed for 27 patients, secondary analyses for all subjects. The primary analysis focused on patients with bacterial LRTI due to culture-confirmed pathogens that had been clinically tested for susceptibility to antimicrobials ($n = 27$ patients, $n = 32$ pathogens) (Fig. 1, Table 1, Additional File 2: Table S2). Of these, 18 patients had respiratory samples sequenced for a prior mNGS study by our group [6]. For secondary analyses, 43 additional patients with clinically adjudicated LRTI and 18 patients with no evidence of LRTI were assessed. In total, the secondary outcome analysis of hospital exposure and AMR gene burden in the respiratory microbiome assessed 70 patients with LRTI and 18 patients with no evidence of LRTI. Assessment of Cas9 targeted Illumina and nanopore sequencing for detecting AMR genes included 10 patients from the primary analysis with culture-confirmed bacterial LRTI.



Procedures

Nucleic acid extraction and Illumina metagenomic sequencing

RNA extraction from mBAL or TA and Illumina metagenomic sequencing were carried out as described previously [6, 14].

Pathogen detection bioinformatics

Detection of respiratory microbes leveraged the ID-Seq pipeline [14] that incorporates the STAR [15] aligner to subtract the human genome (NCBI GRC h38), quality

filtering with PRICESeqfilter [16], and additional filtering to remove non-microbial sequences. The identities of the remaining microbial reads were determined by querying the NCBI nucleotide (NT) and non-redundant protein (NR) databases using GSNAP-L and RAPSEARCH2, respectively [14]. Microbial alignments detected by RNA-seq and DNA-seq were aggregated to the genus level and the sequencing reads comprising each genus were then evaluated for taxonomic assignment at the species level based on species relative abundance. A recently developed rules-based model (RBM) [6] was employed

Table 1 Performance of mNGS for genotypic prediction of antimicrobial susceptibility compared to a reference standard of clinical microbiologic testing. Sensitivity, specificity, and accuracy of DNA+RNA mNGS compared to a reference standard of clinical antimicrobial susceptibility testing based on Clinical & Laboratory Standards Institute (CLSI) minimum inhibitory concentration (MIC) breakpoints. A Gram-positive pathogens. B Gram-negative pathogens. AMR gene(s) detected by mNGS indicated. With respect to genotype-phenotype predictions, squares filled red indicate true positives, squares filled blue indicate true-negatives, squares with purple text = false negatives, squares with orange text are false positives

A Gram-Positive Pathogen	Penicillin/ Ampicillin	Methicillin	Clindamycin/ Erythromycin	Trimethoprim/ Sulfa	Vancomycin	Ceftriaxone	Sens (95% CI)	Spec (95% CI)
	<i>Staphylococcus aureus</i>	BlaZ	TN	ErrA	TN	TN	n/a	100% (16-100%)
<i>Staphylococcus aureus</i>	FN	FN	FN	TN	TN	n/a	0% (0-71%)	100% (16-100%)
<i>Staphylococcus aureus</i>	BlaZ	TN	FN	TN	TN	n/a	50% (1-99%)	100% (29-100%)
<i>Staphylococcus aureus</i>	n/a	MecA	FN	TN	TN	n/a	50% (1-99%)	100% (16-100%)
<i>Staphylococcus aureus</i>	MecA	MecA	MsrA MsrC ErmC	TN	TN	n/a	100% (29-100%)	100% (40-100%)
<i>Staphylococcus aureus</i>	BlaZ	TN	TN	TN	TN	n/a	100% (3-100%)	100% (40-100%)
<i>Staphylococcus aureus</i>	BlaZ	TN	TN	TN	TN	n/a	100% (3-100%)	100% (40-100%)
<i>Staphylococcus aureus</i>	n/a	TN	TN	TN	TN	n/a	n/a	100% (40-100%)
<i>Staphylococcus aureus</i>	BlaZ	TN	MefA	TN	TN	n/a	100% (3-100%)	75% (19-99%)
<i>Staphylococcus aureus</i>	BlaZ	TN	MefA	TN	TN	n/a	100% (3-100%)	75% (19-99%)
<i>Staphylococcus aureus</i>	BlaZ	TN	TN	TN	TN	n/a	100% (3-100%)	100% (40-100%)
<i>Staphylococcus aureus</i>	FN	FN	ErmC	TN	TN	n/a	100% (1-91%)	100% (16-100%)
<i>Enterococcus faecium</i>	Pbp4	n/a	n/a	n/a	VanA	n/a	100% (16-100%)	n/a
<i>Streptococcus pneumoniae</i>	TN	n/a	TN	n/a	TN	TN	n/a	100% (29-100%)
<i>Streptococcus pneumoniae</i>	TN	n/a	TN	n/a	TN	TN	n/a	100% (29-100%)
<i>Streptococcus pneumoniae</i>	TN	n/a	TN	n/a	TN	TN	n/a	100% (29-100%)
<i>Streptococcus pneumoniae</i>	*PBP1a/2x	n/a	TN	n/a	TN	*PBP1a/2x	100% (3-100%)	100% (16-100%)
<i>Streptococcus pneumoniae</i>	TN	n/a	TN	n/a	TN	TN	n/a	100% (29-100%)
<i>Streptococcus pneumoniae</i>	TN	n/a	MefA, MsrD	n/a	TN	TN	n/a	67% (9-99%)
Sens	83% (52-98%)	50% (7-93%)	50% (12-88%)	n/a	100% (3-100%)	100% (3-100%)	Sens	Spec
Spec	100% (48-100%)	100% (63-100%)	73% (43-95%)	100% (74-100%)	100% (81-100%)	100% (48-100%)	70% (47-87%)	95% (85-99%)

B Gram-Negative Pathogen	Ampicillin+ Subactam	Cefazolin	Ceftriaxone/ Ceftazidime	Gentamicin/ Amikacin	Piperacillin- Tazobactam	Trimethoprim/ Sulfa	Ertapenem	Meropenem	Sens (95% CI)	Spec (95% CI)
	<i>Escherichia coli</i>	SST-1 AmpC	SST-1 AmpC	SST-1 AmpC	TN	TN	TN	n/a	n/a	n/a
<i>Klebsiella pneumoniae</i>	SST-1 AmpC	SST-1 AmpC	SST-1 AmpC	TN	TN	TN	n/a	n/a	n/a	50% (12-88%)
<i>Klebsiella pneumoniae</i>	SHV-190, SHV-135	SHV-190, SHV-135	SHV-190, SHV-135	Sat4A, Aph3-III	SHV-190, SHV-135	TN	n/a	n/a	100% (40-100%)	50% (1-99%)
<i>Pseudomonas aeruginosa</i>	OXA-50, OXA-49	n/a	OXA-50, OXA-49	n/a	OXA-50	n/a	n/a	OXA-50	100% (40-100%)	n/a
<i>Klebsiella pneumoniae</i>	SHV-145	SHV-145	SHV-145	TN	SHV-145	TN	n/a	n/a	n/a	33% (0-80%)
<i>Klebsiella aerogenes</i>	AmpC	AmpC	AmpC	TN	n/a	TN	TN	n/a	100% (29-100%)	100% (29-100%)
<i>Enterobacter cloacae</i>	AmpC	AmpC	AmpC	TN	n/a	TN	TN	n/a	100% (29-100%)	100% (29-100%)
<i>Pseudomonas aeruginosa</i>	OXA-50	n/a	OXA-50	Aac6-Ic, Aac3-Ib-9c	OXA-50	n/a	n/a	OXA-50	100% (29-100%)	0% (0-84%)
<i>Stenotrophomonas</i>	n/a	n/a	LI	n/a	n/a	TN	n/a	n/a	100% (3-100%)	100% (3-100%)
<i>Enterobacter cloacae</i>	ACT-MIR	ACT-MIR	ACT-MIR	TN	n/a	TN	TN	n/a	100% (29-100%)	100% (29-100%)
<i>Burkholderia cepaciae</i>	TN	TN	TN	n/a	n/a	n/a	n/a	TN	n/a	100% (40-100%)
<i>Serratia marcescens</i>	SST-1	SST-1	SST-1	Aac6-Ic	TN	TN	n/a	n/a	100% (29-100%)	67% (9-99%)
<i>Serratia marcescens</i>	SST-1	SST-1	SST-1	Aac6-Ic	TN	TN	n/a	n/a	100% (29-100%)	67% (9-99%)
Sens	100% (63-100%)	100% (54-100%)	100% (66-100%)	100% (3-100%)	100% (13-100%)	n/a	n/a	100% (3-100%)	Sens	Spec
Spec	25% (1-81%)	25% (1-81%)	25% (1-81%)	67% (30-93%)	67% (22-96%)	100% (69-100%)	100% (29-100%)	50% (1-99%)	100% (87-100%)	64% (48-78%)

*mutations in PBP1a/2x, Sens Sensitivity, Spec Specificity, TN True negative, FN False negative; n/a phenotypic susceptibility to antibiotic not tested in the clinical laboratory. 95% confidence interval (CI) listed below each sensitivity and specificity value

to differentiate putative pathogens from commensal microbiota.

The RBM leverages previous findings demonstrating that microbial communities in patients with LRTI are typically characterized by one or more dominant pathogens present in high abundance [6, 14]. More specifically, the RBM ranks microbial genera present in a sample by descending abundance (number of taxonomic alignments). The greatest difference between any two sequential taxa is then identified to capture genera present at disproportionately high abundance compared to the rest of the lung microbiota [6, 17].

All genera with an abundance greater than this largest gap threshold are then evaluated at the species level, by identifying the most abundant species within each genus. If the species is present within a previously curated reference index of established respiratory pathogens [6, 17] derived from landmark epidemiologic surveillance studies [18–22], it is selected as a putative pathogen by the RBM. A detailed description of the principles and clinical validation of the RBM has been previously published [6, 17].

Detection of AMR genes

AMR genes present in RNA-seq or DNA-seq data were identified using SRST2 coupled with an expanded version of the ARG-ANNOT database [23] (Additional File 3), and genes with $\geq 5\%$ allele coverage were included in analyses. Because *Streptococcus pneumoniae* is a leading cause of bacterial LRTI [4], we also screened for point mutations in *pbp* genes associated with *Streptococcus* beta lactam resistance using the CARD resistance gene identifier tool and the ‘loose’ setting [24]. Average read depth across each allele, normalized by gene length and total reads (depth per million reads sequenced, dpM), was calculated for each sample.

Assessing performance of genotypic antimicrobial susceptibility prediction

As a reference standard, we used clinical AST results performed in the UCSF Clinical Microbiology Laboratory during each patient’s admission. To calculate sensitivity and specificity, which was done both by microbe and by drug (Table 1), we compared mNGS-based resistance predictions against phenotypic AST determined by the Clinical & Laboratory Standards Institute minimum inhibitory concentration breakpoints [25]. We studied samples from subjects with culture-confirmed bacterial pathogens for which AST was performed. One isolate that only underwent chromogenic beta lactamase screening was excluded. Two subjects (252, 297) with highly polymicrobial cultures of ≥ 4 organisms were also

excluded due to the unclear clinical significance of the isolated microbes. This left a primary analysis cohort of 27 patients and 32 bacterial pathogens (Fig. 1).

We assessed susceptibility to the most common antibiotics used for complicated infections from bacterial pathogens identified in the cohort: *S. aureus*, *S. pneumoniae*, *E. faecium*, *Enterobacteriaceae*, *P. aeruginosa*, and *S. maltophilia*. Initial AMR gene class assignment (beta lactam, aminoglycoside, macrolide/lincosamide/ streptogramin, glycopeptide, trimethoprim/ sulfamethoxazole) was made using ontology in ARG-ANNOT [23] and a more refined AMR phenotype assignment was made based on CARD [24] resistome ontological relationships. In addition to sensitivity and specificity, we assessed very major error (VME; predicted susceptible but phenotypically resistant) and major error (ME; predicted resistant but phenotypically susceptible) rates.

Clinically tested antimicrobials used in mNGS AMR prediction benchmarking

Resistance predictions were made for antibiotics routinely tested in the clinically microbiology laboratory for *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Enterococcus faecium*, *Pseudomonas aeruginosa*, *Stenotrophomonas maltophilia*, and *Enterobacteriaceae*. For *S. aureus* these included penicillin, methicillin, clindamycin or erythromycin, trimethoprim/ sulfamethoxazole (TMP/SMZ), and vancomycin; for *S. pneumoniae*: penicillin, ceftriaxone, and vancomycin; for *E. faecium*: ampicillin and vancomycin; for *Enterobacteriaceae*: ampicillin + sulbactam, ceftazidime, gentamicin, piperacillin-tazobactam, TMP-SMX, ertapenem, and meropenem; for *P. aeruginosa*: ampicillin + sulbactam, ceftazidime, gentamicin, piperacillin-tazobactam, and meropenem; and for *S. maltophilia*: ceftazidime and TMP/SMZ. For some isolates, clinical susceptibility testing for certain antimicrobials was not performed by the clinical laboratory, and thus was unavailable for our analysis.

FLASH Cas9 targeted mNGS for AMR gene detection

FLASH Cas9 targeted Illumina mNGS for AMR gene detection was carried out as described in the original proof of concept study [12]. Briefly, FLASHHit software [26] was first used to design guide RNAs targeting clinically relevant AMR genes derived from the CARD and ResFinder databases, merging exact duplicates [12]. In total, 2226 guide RNAs targeting 381 beta lactam and 111 MLS resistance genes, in addition to the 127 diverse AMR genes from the original FLASH pilot study, were utilized for Cas9 targeted enrichment. Guide RNAs targeted multiple sites on each AMR gene, which in total represented 2226 target sequences (Additional File 4). DNA templates for producing CRISPR RNAs (crRNAs) for each AMR

gene target were synthesized, pooled, transcribed, and purified according to described methods [12].

Ten nanograms of DNA was 5' dephosphorylated using rAPid alkaline phosphatase that was subsequently deactivated with sodium orthovanadate. The dephosphorylated DNA was added to a master mix containing the CRISPR/Cas9 ribonucleoprotein complex and incubated at 37 °C for 2 h. The Cas9 was deactivated with proteinase K and removed with SPRI bead purification. Samples were dA-tailed and then carried forward for Illumina Sequencing according to the NEBNext Ultra II library prep kit (New England Biolabs, Ipswich, MA) protocol according to previously described detailed methods [12]. AMR gene identification was carried out using ARG-ANNOT [23] as for the primary analyses, and genes that were detected at a dpM of >0.1 were assessed for enrichment compared to DNA-seq alone.

Nanopore sequencing

FLASH-enriched DNA libraries were quantified and 200–800 ng of DNA input was used for Nanopore 1D library preparation (protocol SQK-LSK109, Oxford Nanopore, UK). Individual sample libraries were loaded into a single flow cell of a GridION instrument, and sequencing reads were base called in real-time mode in MINKNOW. The SURPIrt pipeline running in -a mode was utilized to identify AMR genes every 100,000–200,000 reads as previously described [27, 28].

Mitigation of background contaminants

To minimize inaccurate taxonomic assignments due to environmental contaminants, we processed negative water controls with each group of samples that underwent nucleic acid extraction, and included these, as well as positive control clinical samples, with each sequencing run. We directly subtracted alignments to those taxa in water control samples detected by both RNA-seq and DNA-seq analyses from the raw reads per million (rpm) values in all samples [6]. To account for selective amplification bias of contaminants in water controls resulting from PCR amplification of metagenomic libraries to a fixed standard concentration across all samples, prior to direct subtraction, we scaled taxa rpms in the water controls to the median percent microbial reads present across all samples as previously described [6]. To address environmental contaminants in AMR gene analyses, resistance alleles detected in water controls at a depth > 1 were excluded.

Study outcomes

The primary outcome was performance of mNGS for predicting phenotypic AST. Secondary

outcomes included the association between hospital exposure and burden of AMR genes in the respiratory microbiome and AMR gene detection using Cas9 targeted enrichment and real-time nanopore sequencing.

Statistical analysis

Statistical significance was defined as *P* less than 0.05, using two-tailed tests of hypotheses. Nonparametric continuous variables were analyzed by Wilcoxon rank-sum.

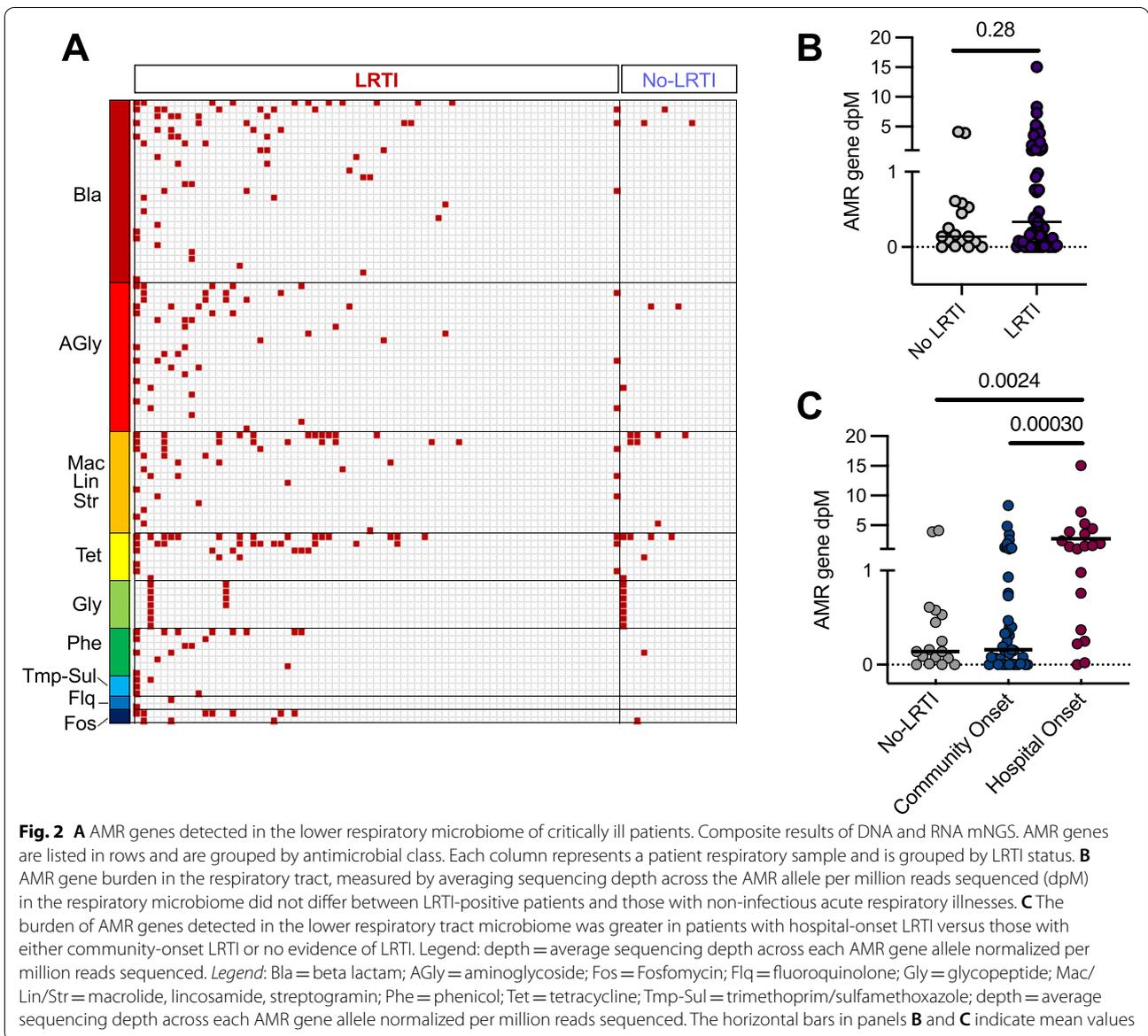
Results

Cohort features

Seventy subjects with LRTI and 18 with no evidence of LRTI were identified based on inclusion and exclusion criteria (Fig. 1, Additional File 1: Table S1). Primary analyses were performed for 27 patients, secondary analyses for all subjects. Clinical AST results were returned a median of 74 h following sample collection (95% confidence interval (CI) 49–115 h, (Additional File 5: Table S3)). Twenty-seven subjects with culture-confirmed bacterial LRTI, representing 32 pathogens with clinical AST data performed on ≥ 2 drugs, were identified and assessed in the primary analysis (Additional File 2: Table S2). For secondary analyses, 43 additional patients with clinically adjudicated LRTI and 18 patients with no evidence of LRTI were assessed.

Metagenomic sequencing, pathogen, and AMR gene detection

A mean of 4.3×10^7 (interquartile range (IQR) $1.9\text{--}4.4 \times 10^7$) DNA-seq reads and 6.9×10^7 (IQR $4.8\text{--}8.3 \times 10^7$) RNA-seq reads were generated from respiratory samples. In the primary AMR analysis group, we used a previously validated [6] metagenomic rules-based model (RBM) to identify bacterial respiratory pathogens that were disproportionately abundant as compared to the rest of the lung microbiome. The RBM identified 26 of 32 (81%) of the culture-confirmed bacterial pathogens from the primary analysis. Four (67%) of the missed pathogens were present in the context of polymicrobial cultures, and one (17%) was identified as a different streptococcal species (Additional File 2: Table S2). A total of 138 and 234 acquired AMR genes were identified by RNA-seq and DNA-seq, respectively (Additional File 6: Table S4). With respect to AMR gene classes, beta lactam resistance genes were most common (81/372 total genes, 35%).



Comparison of mNGS versus phenotypic antimicrobial susceptibility testing

We assessed the performance of mNGS for predicting resistance to clinical guideline-recommended antimicrobials used for complicated gram-negative ($n=8$ drugs) and gram-positive ($n=6$ drugs) infections. AMR genes unrelated to the culture-confirmed bacterial pathogen were identified through the resistome ontology annotations in CARD [24] and excluded from this analysis. Sensitivity and specificity compared to a reference standard of culture-based AST varied by pathogen, drug, patient, and nucleic acid type sequenced (Table 1, Additional File 7: Table S5). For gram-positive pathogens, a combination of DNA-seq and RNA-seq yielded a sensitivity of 70% (CI

47–87%), specificity of 95% (CI 85–99%), and an accuracy of 87% (CI 78–94%) (Table 1). This equated to a VME rate of 30% and a ME rate of 5%. For gram-negative pathogens, a combination of DNA-seq and RNA-seq yielded a sensitivity of 100% (CI 87–100%), specificity of 64% (CI 48–78%), and accuracy of 78% (CI 67–87%) (Table 1). This equated to a VME rate of 0% and a ME rate of 36%.

We also assessed the performance of RNA-seq and DNA-seq performed independently (Table 1, Additional File 7: Table S5). RNA-seq performed with a sensitivity of 52% (CI 31–73%), specificity of 100% (CI 94–100%), and accuracy of 86% (CI 76–93%) for gram-positive pathogens, and a sensitivity of 100% (CI 89–100%), specificity of 64% (CI 48–78%), and accuracy of 79% (CI 68–88%)

for gram-negative pathogens. DNA-seq performed with a sensitivity of 39% (CI 20–61%), specificity of 95% (CI 85–99%), and accuracy of 78% (CI 67–87%) for gram-positive pathogens, and a sensitivity of 58% (CI 39–75%), specificity of 67% (CI 50–80%), and accuracy of 63% (CI 51–74%) for gram-negative pathogens.

In two of seven cases with genotype to phenotype false-positive (ME) predictions, mNGS identified AMR genes unrelated to the culture-confirmed microbe but related to resistant pathogens that would be cultured several days later in the context of ventilator-associated pneumonia (VAP). These included *SST-1* from a patient who developed *Serratia marcescens* VAP (patient 213) 4 days later, and *mecA* from a patient who developed *Staphylococcus aureus* VAP 7 days later (patient 232) (Additional File 6: Table S4).

Association between LRTI positivity, hospital exposure, and AMR genes in the respiratory microbiome

Assessment of the lower respiratory resistome using both DNA and RNA mNGS revealed a diversity of AMR genes in both LRTI-positive and negative patients (Fig. 2A, Additional File 6: Table S4). AMR gene burden did not differ based on LRTI status ($p=0.28$) (Fig. 2B). Subjects with hospital-onset (≥ 48 h after admission) LRTI had a greater burden of AMR genes in their respiratory microbiome compared to those with community-onset LRTI ($p=0.00030$) or those without LRTI ($p=0.0024$) (Fig. 2C).

Targeted enrichment and rapid detection of AMR genes using Cas9 and nanopore sequencing

We validated the utility of a recently described programmable CRISPR/Cas9-based method called FLASH (Finding Low Abundance Sequences by Hybridization) to enrich for low abundance AMR genes [12] by studying 10 patients from the primary analysis (Additional File 8: Table S6). The FLASH+DNA mNGS library prep protocol added 2.5 h to the standard 3-h NEBNext DNA-seq workflow [29] and could enrich detection of AMR genes associated with the culture-confirmed pathogen $>2500\times$ compared to DNA-seq alone (Fig. 3A). In four (40%) patients, FLASH enabled detection of AMR

genes that were associated with the culture-confirmed pathogen and resistance phenotype, but missed by DNA-seq alone, including *mecA* in two patients with methicillin-resistant *Staphylococcus aureus* LRTI. FLASH also resulted in the detection of AMR genes unrelated to the culture-confirmed pathogens in five (50%) of patients (Additional File 8: Table S6).

We subsequently assessed the potential for rapid AMR gene detection using FLASH combined with an Oxford nanopore sequencing platform, which affords real-time data generation (Fig. 3B). All AMR genes identified by Illumina (median sequencing depth of 1.20×10^8 [IQR 5.41×10^7 – 1.36×10^8]) were also identified by FLASH-nanopore sequencing (median sequencing depth of 1.19×10^6 reads [IQR 1.02×10^6 – 1.46×10^6]) (Additional File 9: Table S7). AMR gene targets could be identified within 10 min of real-time nanopore sequencing, suggesting a potential turnaround time of less than 6 h for a single sample (Fig. 3C).

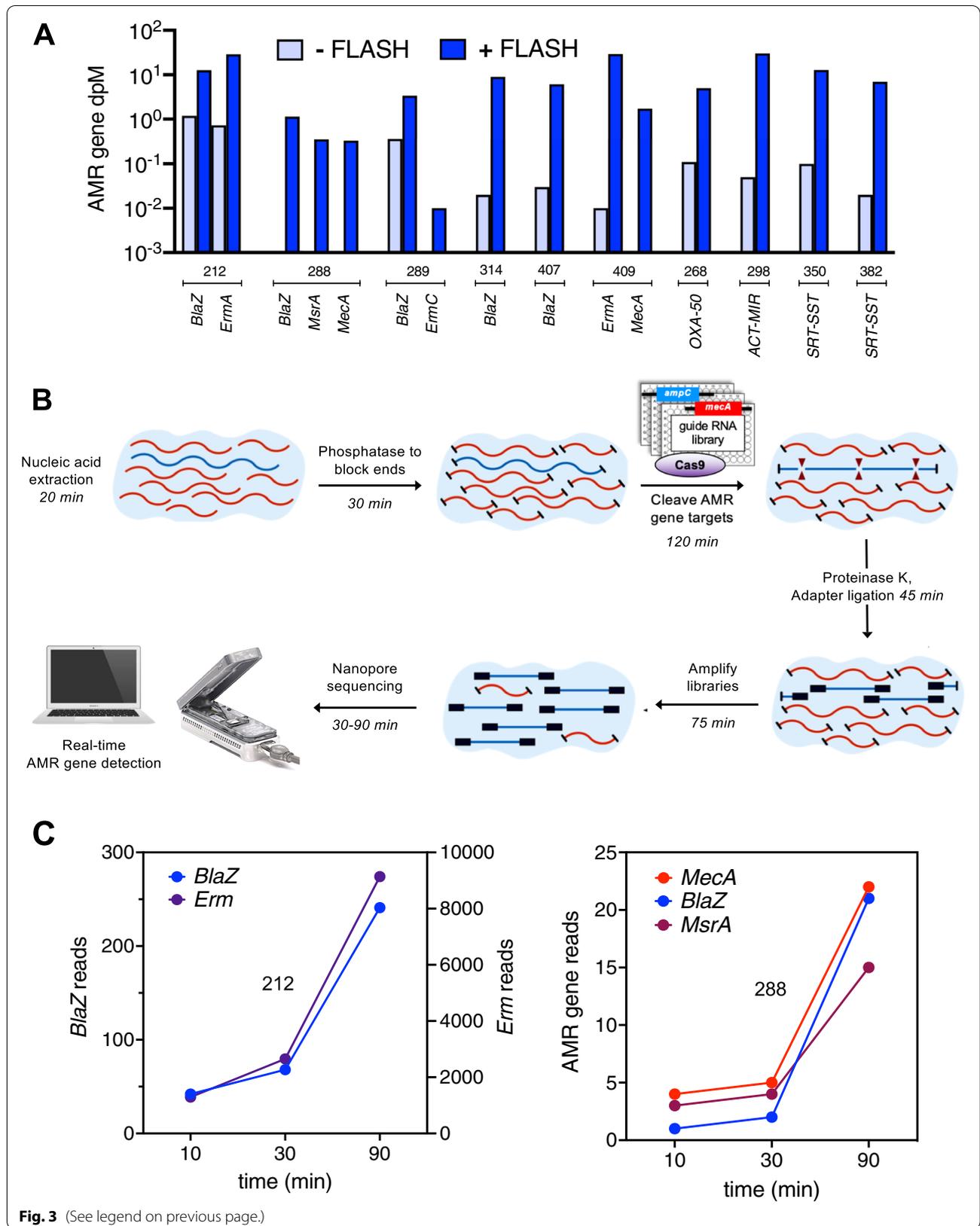
Discussion

Antimicrobial resistance has emerged as one of the most pressing issues facing human health, and effective treatment of complicated infections increasingly necessitates early and accurate assessment of microbial drug resistance. Our study builds on prior respiratory mNGS studies [6, 8–12] by demonstrating that RNA and DNA mNGS can enable culture-independent prediction of AMR in critically ill patients with bacterial LRTI, with variable performance across pathogens and antimicrobials. While false-negative susceptibility predictions for gram-negative pathogens were not observed, we found a VME of 30% for gram-positive pathogens, suggesting mNGS has a role for complementing, rather than replacing, current standard of care culture-based approaches.

Prior work has demonstrated the utility of mNGS in cases of culture-negative LRTI, which represent more than half of all pneumonia cases [4, 6, 8, 10]. Our results suggest that mNGS may also have potential for predicting AST in culture-negative LRTI, where detection of an AMR gene could inform the need for treatment of an occult resistant organism. Further, our findings support recent observations [10] that mNGS may have utility for

(See figure on next page.)

Fig. 3 **A** FLASH (Finding Low Abundance Sequences by Hybridization) CRISPR/Cas9 targeted Illumina sequencing enriched the detection of culture-confirmed bacterial LRTI pathogen AMR alleles by $46\times$ to $>2500\times$ versus DNA-seq alone. **B** Workflow diagram for FLASH targeted enrichment coupled with nanopore sequencing. Time estimates provided for a single sample. **C** Real-time detection of AMR genes by FLASH targeted nanopore sequencing was achieved within 10 min following mNGS library preparation. Data from two representative *Staphylococcus aureus* LRTI cases are highlighted. Case 212 (left panel) highlights a case where detection of *BlaZ* and *MsrA/ErmA* genes correlated with phenotypically determined penicillin and macrolide/lincosamide resistance, respectively. Case 288 (right panel) highlights a case where detection of *MecA*, *BlaZ*, and *MsrA* correlated with phenotypically confirmed methicillin, penicillin, and macrolide resistance, respectively



early identification of future secondary respiratory infections. For instance, patient 232, who was admitted for severe *Klebsiella pneumoniae* LRTI, developed MRSA ventilator-associated pneumonia (VAP) eight days later while undergoing treatment with aztreonam, an antibiotic lacking MRSA coverage. mNGS analysis of a TA specimen obtained 7 days before VAP onset revealed the *mecA* gene, providing an early indication of the subsequent AMR VAP pathogen.

Public health surveillance is essential for pandemic preparedness, understanding trends in AMR, and preventing outbreaks of resistant organisms. Our findings demonstrate the feasibility of mNGS for epidemiological surveillance of AMR and highlight an association between hospital exposure and AMR gene burden in the lower respiratory tract microbiome. Importantly, our analysis included culture-negative LRTI cases and adds to prior literature demonstrating a similar association in cases of culture-confirmed bacterial pneumonia [30].

Nanopore sequencing has proven useful for microbial detection from respiratory samples with high pathogen abundance [8, 10, 31]; however, basecalling accuracy and sensitivity challenges have historically limited its capacity for detecting underrepresented sequences in metagenomic datasets [32]. Targeted enrichment can overcome this through AMR signal amplification, and consistent with this, we found high concordance for AMR gene detection between nanopore and Illumina samples that underwent FLASH Cas9 targeted enrichment. Our results suggest that this method could also potentially augment AMR gene detection and resistance prediction if coupled with established nanopore mNGS workflows for detecting respiratory pathogens [8–10].

Rapid detection of AMR is essential in critically ill patients with severe bacterial infections given that time to appropriate antimicrobials correlates with mortality [33]. FLASH adds 2.5 h to standard DNA library preparation workflow and, when coupled with real-time detection of AMR genes from mNGS libraries, could potentially allow for sample to answer in under 6 h, a significant time savings compared to the ≥ 24 h required for Illumina protocols [34]. In our cohort, clinical AST required an average of 74 h, suggesting that Cas9-targeted nanopore sequencing may be a promising future approach for more rapidly identifying patients with resistant infections. We found that FLASH also enriched for AMR genes unrelated to the culture-confirmed pathogen, presumably derived from the lung microbiome. Improved methods for annotating the species-specificity of detected AMR genes may help address this issue, which otherwise could lead to increased ME due to false-positive results.

Strengths of this study include that it is the largest to assess the performance of mNGS AMR prediction in

LRTI and that it systematically assessed performance for multiple classes of antimicrobials against a clinical reference standard. Further, we provide the first culture-independent assessment of healthcare exposure and resistance gene burden in the respiratory tract, and a novel demonstration of Cas9 targeted enrichment coupled with rapid nanopore sequencing. Our study also has several limitations, including sample size, spectrum of antimicrobial classes assessed, spectrum of pathogens assessed, and the need for independent validation of findings. Future work in a larger, prospective cohort with a greater diversity of bacterial pathogens and resistance mechanisms can address these limitations. In addition, a randomized clinical trial will be needed to assess the potential impact of mNGS on time to appropriate antimicrobial treatment, antimicrobial stewardship, and LRTI outcomes. Lastly, genotype to phenotype prediction remains imperfect, even for cultured isolates [7, 35]. As with other molecular testing modalities for AMR, this should be recognized when considering the utility and clinical applicability of mNGS for AMR phenotypic prediction.

Conclusions

In summary, we characterize the utility of mNGS for predicting AMR in bacterial LRTI and demonstrate proof of concept for both epidemiological AMR surveillance and rapid resistance gene detection using Cas9 and nanopore sequencing.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-022-01072-4>.

Additional file 1: Table S1. Clinical and demographic features of cohort.

Additional file 2: Table S2. mNGS detection of respiratory pathogens compared to a reference standard of clinical microbiological testing.

Additional file 3: Dataset 1. AMR gene database utilized for analyses.

Additional file 4: Dataset 2. AMR genes and guide RNAs for FLASH targeted mNGS. AMR genes and guide RNAs for FLASH targeted mNGS. A) AMR genes targeted by FLASH. B) AMR alleles and guide RNA targets. Legend: Bla = beta lactam; MLS = macrolide, Tmp/Sul = trimethoprim/sulfamethoxazole.

Additional file 5: Table S3. Time to return of clinical antimicrobial susceptibility testing results for culture-confirmed bacterial pathogens.

Additional file 6: Table S4. AMR genes detected by mNGS.

Additional file 7: Table S5. Performance of RNA-seq, DNA-seq and combined RNA-seq + DNA-seq for AMR prediction compared to clinical antimicrobial susceptibility testing.

Additional file 8: Table S6. Pathogen AMR genes detected by FLASH-mNGS.

Additional file 9: Table S7. AMR genes detected by FLASH coupled with Nanopore mNGS.

Acknowledgements

Not applicable.

Authors' contributions

PHS, XD: collection and assembly of data, data analysis and interpretation, manuscript editing. MA, KM, SC, MF, NN: collection and assembly of data, manuscript editing. RG, TD: collection and assembly of data. EC, AL.: methods development and manuscript editing. AM, KK: data analysis and interpretation. SD, SM: manuscript editing and data interpretation. CCa, CCh, JD: conception and design, manuscript editing, funding. CL: conception and design, data analysis and interpretation, manuscript writing, funding. All authors read and approved the final manuscript.

Funding

NIH R35 HL140026 (CSC), NHLBI K23HL138461-01A1 (CL), Chan Zuckerberg Biohub (CL, JLD). The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Availability of data and materials

Raw microbial sequences are available via NCBI BioProject PRJNA450137 [36], <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA450137>, and BioProject PRJNA635133 [37], <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA635133>. FLASHit [26] and SURPIrt [28] have been deposited on Github and are available for download at: <https://github.com/czbiohub/flash> and <https://github.com/chulab/SURPI-plus-dist>, respectively.

Declarations

Ethics approval and consent to participate

Patients were enrolled via a previously described observational cohort study [6, 38] which was approved by the UCSF Institutional Review Board (IRB), which granted a waiver of initial consent for respiratory sampling. Informed consent was subsequently obtained from patients or their surrogates for continued study participation, following the previously detailed IRB protocol #17–24056 [38]. This research conformed to the principles of the Helsinki Declaration.

Consent for publication

Consent to publish patient-related information was obtained as part of informed consent for study participation as described above.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Division of Infectious Diseases, Department of Medicine, University of California, San Francisco, San Francisco, CA, USA. ²Chan Zuckerberg Biohub, San Francisco, CA, USA. ³Department of Laboratory Medicine, University of California, San Francisco, CA, USA. ⁴Department of Microbiology and Immunology, University of California, San Francisco, CA, USA. ⁵Division of Pulmonary, Critical Care, Allergy and Sleep Medicine, Department of Medicine, University of California, San Francisco, CA, USA. ⁶Chan Zuckerberg Initiative, San Francisco, CA, USA. ⁷Department of Biochemistry and Biophysics, University of California, San Francisco, CA, USA.

Received: 6 February 2022 Accepted: 15 June 2022

Published online: 12 July 2022

References

- O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. London, UK May 16, 2016. 2016; published online May 16. <http://amr-review.org/Publications>.
- Zaas AK, Garner BH, Tsalik EL, Burke T, Woods CW, Ginsburg GS. The current epidemiology and clinical decisions surrounding acute respiratory infections. *Trends Mol Med*. 2014;20:579–88.
- van der Eerden MM, Vlaspolder F, de Graaff CS, et al. Comparison between pathogen directed antibiotic treatment and empirical broad spectrum antibiotic treatment in patients with community acquired pneumonia: a prospective randomised study. *Thorax*. 2005;60:672–8.
- Jain S, Self WH, Wunderink RG, et al. Community-acquired pneumonia requiring hospitalization among U.S. adults. *N Engl J Med*. 2015;373:415–27.
- Leffler DA, Lamont JT. Clostridium difficile infection. *N Engl J Med*. 2015;372:1539–48.
- Langelier C, Kalantar KL, Moazed F, et al. Integrating host response and unbiased microbe detection for lower respiratory tract infection diagnosis in critically ill adults. *Proc Natl Acad Sci USA*. 2018;115(52):E12353–62 (201809700).
- Mahfouz N, Ferreira I, Beisken S, von Haeseler A, Posch AE. Large-scale assessment of antimicrobial resistance marker databases for genetic phenotype prediction: a systematic review. *J Antimicrob Chemother*. 2020;75:3099–108.
- Charalampous T, Kay GL, Richardson H, et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat Biotechnol*. 2019;37:783–92.
- Yang L, Haidar G, Zia H, et al. Metagenomic identification of severe pneumonia pathogens in mechanically-ventilated patients: a feasibility and clinical validity study. *Respir Res*. 2019;20:265.
- Charalampous T, Alcolea-Medina A, Snell LB, et al. Evaluating the potential for respiratory metagenomics to improve treatment of secondary infection and detection of nosocomial transmission on expanded COVID-19 intensive care units. *Genome Med*. 2021;13:182.
- Chao L, Li J, Zhang Y, Pu H, Yan X. Application of next generation sequencing-based rapid detection platform for microbiological diagnosis and drug resistance prediction in acute lower respiratory infection. *Ann Transl Med*. 2020;8:1644.
- Quan J, Langelier C, Kuchta A, et al. FLASH: a next-generation CRISPR diagnostic for multiplexed detection of antimicrobial resistance sequences. *Nucleic Acids Res* 2019; published online May 22. <https://doi.org/10.1093/nar/gkz418>.
- United States Centers for Disease Control and Prevention. CDC/NHNS surveillance definitions for specific types of infections. 2021; published online Jan. https://www.cdc.gov/nhns/pdfs/psmanual/psmanual_current.pdf.
- Kalantar KL, Carvalho T, de Bourcy CFA, et al. IDseq—An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *Gigascience* 2020; 9. <https://doi.org/10.1093/gigascience/giaa111>.
- Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
- Ruby JG, Bellare P, Derisi JL. PRICE: software for the targeted assembly of components of (Meta) genomic sequence data. *G3 (Bethesda)*. 2013;3:865–80.
- Tsitsiklis A, Osborne CM, Kamm J, et al. Lower respiratory tract infections in children requiring mechanical ventilation: a multicentre prospective surveillance study incorporating airway metagenomics. *Lancet Microbe*. 2022;3:e284–93.
- Jain S, Self WH, Wunderink RG, et al. Community-acquired pneumonia requiring hospitalization among U.S. adults. *N Engl J Med*. 2015;373:415–27.
- Infection in Organ Transplantation - Fishman - 2017 - American Journal of Transplantation - Wiley Online Library. <https://doi.org/10.1111/ajt.14208> (Accessed 21 Jul 2021).
- Magill SS, Edwards JR, Bamberg W, et al. Multistate point-prevalence survey of health care-associated infections. *N Engl J Med*. 2014;370:1198–208.
- Kalil AC, Metersky ML, Klompas M, et al. Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 Clinical Practice Guidelines by the Infectious Diseases Society of America and the American Thoracic Society. *Clin Infect Dis*. 2016;63:e61–111.
- Mandell LA, Wunderink RG, Anzueto A, et al. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin Infect Dis*. 2007;44(Suppl 2):S27–72.
- Inouye M, Dashnow H, Raven L-A, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med*. 2014;6:90.

24. Jia B, Raphenya AR, Alcock B, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 2017;45:D566–73.
25. CLSI. CLSI M100-ED31:2021 Performance Standards for Antimicrobial Susceptibility Testing, 31st Edition. 2021; published online March. <http://em100.edaptivedocs.net/dashboard.aspx>.
26. Quan J, Langelier CR, Batson J, Crawford ED. FLASHHit. Github. <https://github.com/czbiohub/flash> 2020.
27. Deng X, Achari A, Federman S, et al. Metagenomic sequencing with spiked primer enrichment for viral diagnostics and genomic surveillance. *Nature Microbiology* 2020; published online Jan 13. <https://doi.org/10.1038/s41564-019-0637-9>.
28. Chiu CY, Stryke D. SURPI+, a rapid computational pipeline for comprehensive identification of pathogens from clinical metagenomic sequence data. Github. <https://github.com/chiulab/SURPI-plus-dist> 2019.
29. New England Biolabs. Improved library preparation with the NEBNext® UltraTM II DNA Library Prep Kit for Illumina. <https://www.neb-online.de/wp-content/uploads/2020/03/Improved-library-preparation-with-the-NEBNext-Ultra-II-DNA-Library-Prep-Kit-for-Illumina-E7645.pdf> (Accessed 9 Dec 2021).
30. Kalil AC, Metersky ML, Klompas M, et al. Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 Clinical Practice Guidelines by the Infectious Diseases Society of America and the American Thoracic Society. *Clin Infect Dis.* 2016;63:e61–111.
31. Greninger AL, Naccache SN, Federman S, et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* 2015;7:99. <https://doi.org/10.1186/s13073-015-0220-9>.
32. Rang FJ, Kloosterman WP, de Ridder J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol.* 2018;19:90. <https://doi.org/10.1186/s13059-018-1462-9>.
33. Lee C-C, Lee C-H, Yang C-Y, Hsieh C-C, Tang H-J, Ko W-C. Beneficial effects of early empirical administration of appropriate antimicrobials on survival and defervescence in adults with community-onset bacteremia. *Crit Care.* 2019;23:363.
34. Illumina, Inc. Run time estimates for each sequencing step on the Illumina sequencing platforms. 2020; published online May 15. <https://emea.support.illumina.com/bulletins/2017/02/run-time-estimates-for-each-sequencing-step-on-illumina-sequenci.html> (Accessed 9 Dec 2021).
35. Yee R, Dien Bard J, Simmer PJ. The genotype-to-phenotype dilemma: how should laboratories approach discordant susceptibility results? *J Clin Microbiol.* 2021;59:e00138–e220.
36. Langelier CR, Kalantar KL. Combined host and microbe NGS for lower respiratory tract infection diagnosis in critically ill adults microbial sequence reads. BioProject PRJNA450137, Sequence Read Archive. 2018. <https://www.ncbi.nlm.nih.gov/bioproject/?term=%20PRJNA450137>.
37. Langelier CR, Hayakawa Serpa P. CAS9-mNGS detects antimicrobial resistant pathogens. BioProject PRJNA635133, Sequence Read Archive. 2020. <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA635133>.
38. Sarma A, Christenson SA, Byrne A, et al. Tracheal aspirate RNA sequencing identifies distinct immunological features of COVID-19 ARDS. *Nat Commun.* 2021;12:5152.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

