

1 **Metagenomic identification of severe pneumonia pathogens with rapid Nanopore sequencing in** 2 **mechanically-ventilated patients.**

3
4 Libing Yang, MD^{1,2}; Ghady Haidar, MD³; Haris Zia, MD⁴; Rachel Nettles, MS¹; Shulin Qin, MD, PhD^{1,2}; Xiaohong
5 Wang, MS^{1,2}; Faraaz Shah, MD^{2,5}; Sarah F. Rapport, BS/MPH²; Themoula Charalampous, PhD⁶; Barbara Methé,
6 PhD^{1,2}; Adam Fitch, MS¹; Alison Morris, MD, MS^{1,2,7}; Bryan J. McVerry, MD^{1,2}; Justin O'Grady, PhD^{6,8}, Georgios
7 D. Kitsios, MD, PhD^{1,2}

8
9 ¹Center for Medicine and the Microbiome, University of Pittsburgh;

10 ²Division of Pulmonary, Allergy and Critical Care Medicine, Department of Medicine, University of Pittsburgh
11 School of Medicine and University of Pittsburgh Medical Center, Pittsburgh, PA, USA;

12 ³Division of Infectious Diseases, University of Pittsburgh Medical Center and University of Pittsburgh School of
13 Medicine, Pittsburgh, PA, USA;

14 ⁴Internal Medicine Residency Program, University of Pittsburgh Medical Center McKeesport;

15 ⁵Veterans Affairs Pittsburgh Healthcare System

16 ⁶Bob Champion Research and Educational Building, University of East Anglia, Norwich Research Park, Norwich,
17 UK;

18 ⁷Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, USA;

19 ⁸Quadram Institute Bioscience and University of East Anglia.

20 21 **Author contributions**

22 Conception and design: LY, AM, BJM, JOG, GDK

23 Acquisition, analysis or interpretation of data: LY, RN, HZ, SQ, XW, FS, SFR, TC, BM, AF, GH, AM, BJM, JOG,
24 GDK

25 Drafting of work and/or revising for important intellectual content: LY, AM, BJM, GDK

26 Final approval of version to be published; agreement to be accountable for all aspects of the work in ensuring
27 that questions related to the accuracy or integrity of any part of the work are appropriately investigated and
28 resolved: LY, RN, HZ, SQ, XW, FS, SFR, TC, BM, AF, GH, AM, BJM, JOG, GDK

29

30 **Funding support:**

31 Funding support: National Institutes of Health [K23 HL139987 (GDK); U01 HL098962 (AM); P01 HL114453
32 (BJM); R01 HL097376 (BJM); K24 HL123342 (AM); K23 GM122069 (FS)]. This paper presents independent
33 research funded by the National Institute for Health Research (NIHR) under its Program Grants for Applied
34 Research Program (reference no. RP-PG-0514-20018, JOG.), the UK Antimicrobial Resistance Cross Council
35 Initiative (no. MR/N013956/1, JOG), Rosetrees Trust (no. A749, JOG) and the Biotechnology and Biological
36 Sciences Research Council (BBSRC) Institute Strategic Programme Microbes in the Food Chain BB/R012504/1
37 and its constituent projects BBS/E/F/000PR10348 and BBS/E/F/000PR10349 (JOG).

38

39 **Conflicts of Interest:**

40 Dr. Bryan J. McVerry is a consultant for Vapotherm, Inc. Dr. Georgios Kitsios receives research funding from
41 Karius, Inc. Dr. Justin O'Grady receives (or received) research funding and consumable support from Oxford
42 Nanopore Technologies (ONT) and financial support for attending conferences and for speaking at ONT
43 headquarters. The other authors have no conflicts of interest to declare.

44

45

46

47

48 **Abstract**

49 **Background:** Metagenomic sequencing of respiratory microbial communities for etiologic pathogen identification
50 in pneumonia may help overcome the limitations of current culture-based methods. We examined the feasibility
51 and clinical validity of rapid-turnaround metagenomics with Nanopore™ sequencing of respiratory samples for
52 severe pneumonia diagnosis.

53
54 **Methods and Findings:** We conducted a case-control study of mechanically-ventilated patients with pneumonia
55 (nine culture-positive and five culture-negative) and without pneumonia (eight controls). We collected
56 endotracheal aspirate samples (ETAs) and applied a microbial DNA enrichment method prior to performing
57 metagenomic sequencing with the Oxford Nanopore MinION device. We compared Nanopore results against
58 clinical microbiologic cultures and bacterial 16S rRNA gene sequencing. In nine culture-positive cases, Nanopore
59 revealed communities with low alpha diversity and high abundance of the bacterial (n=8) or fungal (n=1) species
60 isolated by clinical cultures. In four culture-positive cases with resistant organisms, Nanopore detected antibiotic
61 resistance genes corresponding to the phenotypic resistance identified by clinical antibiograms. In culture-
62 negative pneumonia, Nanopore revealed probable bacterial pathogens in 1/5 cases and airway colonization by
63 *Candida* species in 3/5 cases. In controls, Nanopore showed high abundance of oral bacteria in 5/8 subjects,
64 and identified colonizing respiratory pathogens in the three other subjects. Nanopore and 16S sequencing
65 showed excellent concordance for the most abundant bacterial taxa.

66
67 **Conclusion:** We demonstrated technical feasibility and proof-of-concept clinical validity of Nanopore
68 metagenomics for severe pneumonia diagnosis, with striking concordance with positive microbiologic cultures
69 and clinically actionable information offered from the sequencing profiles of culture-negative samples.
70 Prospective studies with real-time metagenomics are warranted to examine the impact on antimicrobial decision-
71 making and clinical outcomes.

74 **Introduction**

75 Pneumonia is a primary cause of morbidity and mortality among adults, leading to more than one million
76 hospitalizations every year and high rates of intensive care unit (ICU) admission in the US (1). The mainstay of
77 pneumonia management is early and appropriate antimicrobial therapy targeting the causative pathogens,
78 balanced with preventing antibiotic overuse and emergence of resistance (2). Thus, timely and accurate
79 identification of causal pathogens is imperative yet remains challenging due to reliance on culture-based
80 methods with low sensitivity and long turnaround times (48~72 hours to actionable results (3)). Recently
81 developed rapid polymerase-chain reaction (PCR) tests represent a significant advancement in the field (4), but
82 these tests can only detect the presence/absence of selected panels of pathogens, and thus are not
83 comprehensive enough in breadth or resolution. Culture-independent methods using next-generation
84 sequencing of microbial communities may help overcome the limitations of current diagnostic testing (5–7).

85 Our group and others have provided proof-of-concept evidence that sequencing of the bacterial 16S
86 rRNA gene (16S sequencing) in clinical respiratory specimens can provide diagnostic insights beyond standard
87 microbiologic cultures (5,8–10). Nevertheless, standard 16S sequencing is not clinically applicable due to limited
88 resolution (providing only genus-level bacterial identification) and lengthy sample processing, library preparation
89 and data acquisition timelines (11). The advent of Nanopore metagenomic sequencing (Oxford Nanopore
90 Technologies [ONT], UK) has offered unprecedented capacities for real-time, detailed profiling of microbial
91 communities at species level (including viruses and fungi) (12–15). With recent technical improvements to
92 overcome the high amounts of contaminating human DNA in clinical respiratory samples (16), Nanopore
93 metagenomics may allow for the development of a novel diagnostic approach for pneumonia.

94 In this study, we sought to evaluate the technical feasibility and clinical validity of Nanopore metagenomic
95 sequencing for etiologic diagnosis of severe pneumonia in mechanically-ventilated patients in the ICU.

97 **Materials and Methods**

98 Detailed methods are provided in the Supplement.

99 **Study design and Participants**

100 From June 2018 – March 2019, we carried out a nested case-control study from an ongoing registry
101 enrolling mechanically-ventilated adult patients with acute respiratory failure in the Medical Intensive Care Unit

(MICU) at the University of Pittsburgh Medical Center (UPMC) (5,17). Exclusion criteria included inability to obtain informed consent, presence of tracheostomy, or mechanical ventilation for more than 72 hours prior to enrollment.

We diagnosed clinical pneumonia based on consensus committee review of clinical, radiographic, and microbiologic data per established criteria (18). We selected 14 subjects with a clinical diagnosis of pneumonia (nine with culture-confirmed diagnosis [culture-positive pneumonia group] and five with negative cultures [culture-negative pneumonia group]). Culture-positivity was deemed when a probable respiratory pathogen was isolated in clinical microbiologic cultures of respiratory specimens obtained at the discretion of treating physicians (sputum, endotracheal aspirate [ETA], or bronchoalveolar lavage fluid [BALF]). In culture-positive cases, antibiotic susceptibility testing was done as per standard practice at UPMC's clinical microbiology laboratory, and results were interpreted based on Clinical and Laboratory Standards Institute criteria (19). We also included eight subjects (controls) who did not have evidence of lower respiratory tract infection and were intubated either for airway protection (n=5) or respiratory failure from cardiogenic pulmonary edema (n=3). From enrolled subjects, we collected ETAs for research purposes (sequencing) within the first 48hr from intubation. In two subjects, we utilized ETAs obtained on the fifth day post-intubation (instead of their baseline samples) when there was clinical suspicion of ventilator-associated pneumonia (VAP) (Case 10 and 15). From plasma samples taken at the same time with ETAs, we measured plasma procalcitonin levels (17). We recorded demographic, physiological, and laboratory variables at the time of sample acquisition, from which we calculated clinical pulmonary infection scores (CPIS) (20), and reviewed the antimicrobial therapies administered for the first 10 days from intubation.

This study was approved by the University of Pittsburgh Institutional Review Board (Protocol PRO10110387). Written informed consent was provided by all participants or their surrogates in accordance with the Declaration of Helsinki.

Microbial DNA sequencing approaches

We focused our sequencing approach on DNA-based organisms (i.e. excluding RNA viruses) and aimed to perform agnostic profiling for microbes (bacteria and fungi) present in the ETAs obtained from the patients in the ICU. However, metagenomic microbial DNA sequencing in clinical respiratory samples is technically challenging because of the high amounts of contaminating human DNA that can overwhelm the sequencing

130 output (ratio of human:microbial DNA >99:1 (21). Therefore, we applied a human DNA depletion step in ETA
131 samples that utilized a detergent-based (saponin) method for selective lysis of human cells and digestion of
132 human DNA with nuclease, as recently described (16). We extracted genomic DNA with the DNeasy Powersoil
133 Kit (Qiagen, Germantown, MD) and assessed the efficiency of human DNA depletion by comparing quantitative
134 PCR (qPCR) cycle threshold (Ct) of a human gene (Glyceraldehyde 3-phosphate dehydrogenase - GAPDH) and
135 the bacterial 16S rRNA gene (V3-V4 region) (22) between samples subjected to depletion vs. not (depleted vs.
136 undepleted samples).

137 From extracted DNA in depleted samples, we prepared metagenomic sequencing libraries with a Rapid
138 PCR Barcoding Kit (SQK-RPB004) and then ran on the MinION device (Oxford Nanopore Technologies (ONT),
139 UK, UK) for five hours. We basecalled the output (i.e. converted the sequencing device output into nucleic acid
140 base sequences) with the Guppy software and used the ONT platform, EPI2ME, for quality control, species
141 identification [What's In My Pot (WIMP) pipeline] and antimicrobial resistance gene analyses [ARMA workflow].
142 Samples that generated fewer than 300 high-quality microbial reads were excluded from further analyses. As an
143 internal quality control for the reliability and reproducibility of Nanopore sequencing, we performed sequencing
144 on two samples with extracted DNA from a mock microbial community with known composition (ZymoBIOMICS
145 Microbial Community Standard) and compared derived vs. expected abundance of microbial species.

146 To further validate the results of Nanopore sequencing for bacterial DNA, we performed standard 16S
147 rRNA gene (V4 region) PCR amplification and sequencing on the Illumina MiSeq Platform as a reference method
148 for bacterial DNA sequencing (23). We processed 16S sequences using an in-house pipeline developed by the
149 University of Pittsburgh Center for Medicine and the Microbiome (CMM) (24–29). Samples that generated fewer
150 than 100 bacterial reads were excluded from further analyses.

151 **Ecological and statistical analyses**

152 From sequencing reads obtained from Nanopore and 16S sequencing, we calculated alpha diversity by
153 Shannon index, performed permutational multivariate analysis of variance (PERMANOVA) testing to assess
154 compositional differences between sample types, and visualized compositional dissimilarities between samples
155 with the non-metric multidimensional scaling (NMDS) method using the Bray-Curtis index. All analyses were
156 performed with the R *vegan* package (30).

157

158 Results

159 Cohort description

160 We enrolled 22 mechanical-ventilated patients with acute respiratory failure: nine with retrospective
 161 consensus diagnosis of culture-positive pneumonia, five with culture-negative pneumonia, and eight controls.
 162 Clinical characteristics and outcomes for the three groups are shown in Table 1. Cases with culture-positive
 163 pneumonia had significantly higher CPIS and a trend for higher procalcitonin levels compared to controls (Table
 164 1, Fig S1). At the time of enrollment, empiric antibiotics had been prescribed for all 14 patients with clinical
 165 diagnosis of pneumonia, as well as for 5/8 of control patients (Table 1, Fig S3).

167 **Table 1: Characteristics of enrolled patients.** Continuous variables are presented as medians (with
 168 interquartile ranges), and categorical variables are presented as N (%).

	Culture-Positive Pneumonia	Culture-Negative Pneumonia	Controls
N	9	5	8
Age, median [IQR], yrs	58.3 [55.2, 62.6]	55.8 [45.7, 62.8]	61.2 [51.9, 67.4]
Male, N (%)	5 (55.6)	2 (40.0)	6 (75.0)
BMI, median [IQR]	24.0 [21.5, 34.6]	31.2 [25.6, 32.9]	28.1 [25.1, 36.1]
SOFA Score, median [IQR]*	6.0 [4.0, 6.0]	7.0 [6.0, 7.0]	5.0 [4.0, 8.0]
PaO₂:FIO₂ ratio, median [IQR], mmHg	158.0 [137.0, 275.0]	150.0 [121.0, 208.0]	221.5 [205.0, 319.5]
Heart rate (median [IQR]), beats per minute	107.0 [92.0, 117.0]	83.0 [82.0, 88.0]	81.5 [73.8, 85.5]
SBP (median [IQR])	125.0 [102.0, 141.0]	118.0 [117.0, 127.0]	105.0 [97.8, 117.5]
WBC, median [IQR], x 10⁹ per liter (L)	10.0 [7.4, 16.8]	4.6 [3.5, 8.3]	5.4 [4.5, 11.6]
Platelets, median [IQR], x 10⁹ per liter (L)	190.0 [169.0, 281.0]	155.0 [136.0, 210.0]	141.0 [72.2, 171.0]
Creatinine, median [IQR], mg/dL	1.3 [1.1, 1.5]	1.5 [1.5, 2.0]	1.0 [0.8, 1.4]
Respiratory Rate, median [IQR], 1/min	22.0 [21.0, 24.0]	21.0 [20.0, 24.0]	17.0 [15.5, 17.2]
PEEP, median [IQR], cm	8.0 [5.0, 8.0]	5.0 [5.0, 8.0]	5.0 [5.0, 5.8]
Tidal Volume (per kg of PBW), (median [IQR]), ml/kg	6.8 [6.2, 8.4]	6.2 [6.1, 6.6]	6.3 [6.0, 7.1]
Plateau Pressure, median [IQR], cm	20.0 [13.0, 23.5]	25.0 [21.0, 29.0]	16.0 [13.0, 21.5]
Ventilator free days, median [IQR], days	12.0 [0.0, 23.0]	17.0 [6.0, 23.0]	24.5 [24.0, 26.0]

ICU Length of Stay, median [IQR], days	8.0 [5.0, 18.0]	11.0 [6.0, 12.0]	4.5 [3.8, 5.0]
Acute Kidney Injury, N (%)	8 (88.9)	4 (80.0)	2 (25.0)
30 Day Mortality, N (%)	3 (33.3)	1 (20.0)	1 (12.5)
On antibiotic therapy, N (%)	9 (100)	5 (100)	5 (62.5)
Antibiotic days, median [IQR], days	12.0 [10.0, 24.0]	21.0 [20.0, 24.0]	3.5 [0.0, 7.5]
CPIS, median [IQR],	8.0 [7.0, 9.0]	6.0 [5.0, 7.0]	5.0 [4.0, 5.2]
Procalcitonin, median [IQR], pg/μl	2783.0 [1049.5, 4330.1]	4866.0 [94.4, 4965.1]	353.5 [250.4, 1531.6]

169

170

171

172

173

174

175

176

177

178

Abbreviations: IQR: interquartile range; BMI: body mass index; SOFA: sequential organ failure assessment; PaO₂: partial pressure of arterial oxygen; FiO₂: fractional inhaled concentration of oxygen; SBP: systolic blood pressure; WBC: white blood cell count; PEEP: positive end-expiratory pressure; PBW: predicted body weight; ICU: intensive care unit; CPIS: clinical pulmonary infection score.

* SOFA score calculation does not include the neurologic component of SOFA score because all patients were intubated and receiving sedative medications, impairing our ability to perform assessment of the Glasgow Coma Scale in a consistent and reproducible fashion.

178

Technical feasibility of Nanopore sequencing in clinical samples

179

180

181

182

183

184

185

186

187

188

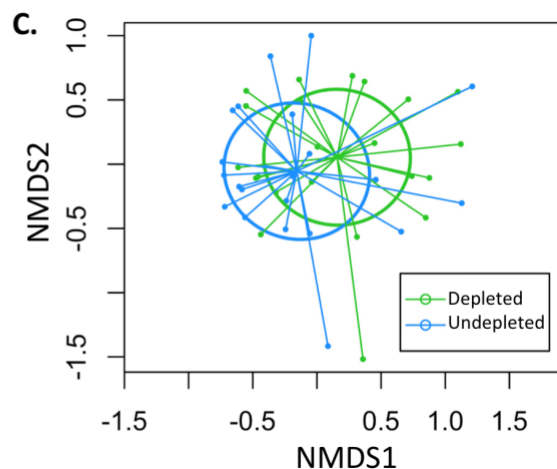
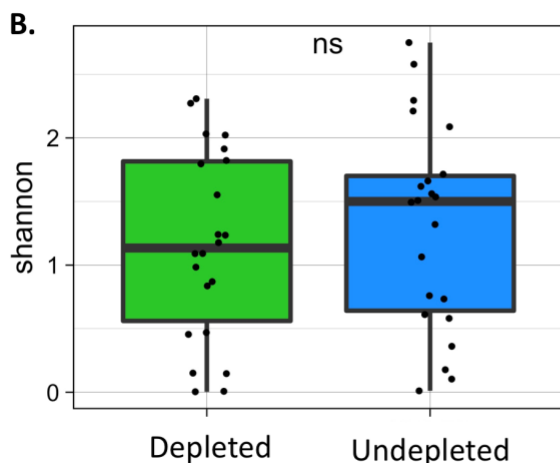
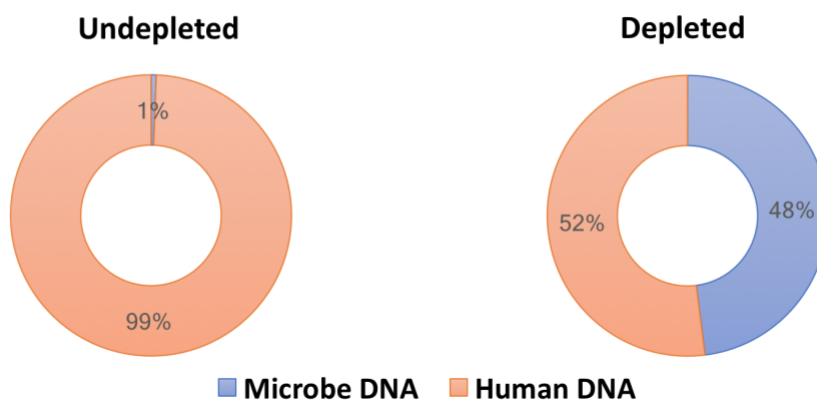
189

Pre-processing of the ETA samples with the saponin-based human DNA depletion protocol resulted in relative enrichment of bacterial DNA by an average of 1260-fold (Fig S2). This microbial enrichment step allowed for generation of sufficient numbers of microbial reads by Nanopore sequencing in depleted samples (median 6682 reads, average proportion 48% of total reads), whereas in undepleted samples the sequencing output was overwhelmed by human DNA (only 1% of reads were of microbial origin) and effectively was unusable (Fig 1A). Importantly, the depletion protocol did not appear to alter the underlying bacterial communities, because ecological analyses of depleted and undepleted samples by 16S rRNA gene sequencing demonstrated no significant differences in alpha or beta diversity (Fig 1B, C).

Fig 1. Saponin-based human DNA depletion effectively removed human DNA without changing bacterial community structure. (A) Before human DNA depletion, 1% of Nanopore reads were of microbial origin;

190 following human DNA depletion, 48% of Nanopore reads were of microbial origin microbes. (B) There was no
191 significant difference in alpha diversity of bacterial communities between depleted and undepleted samples
192 assessed by 16S rRNA gene sequencing. (C) Non-metric multidimensional scaling (NMDS) plot of the Bray-
193 Curtis dissimilarity index between depleted and undepleted samples based on 16S rRNA gene sequencing.
194 Depleted samples were compositionally similar to undepleted samples (PERMANOVA, p -value=0.17).

A. Proportion of microbe reads vs. human reads by Nanopore



195

196

197 Analytical validity of Nanopore sequencing

198 Nanopore-derived bacterial communities showed striking similarity with both mock communities of
199 extracted DNA (table S1) as well as 16S-derived community profiles for bacteria from clinical samples (Fig S3),
200 underscoring the analytical validity of Nanopore results for use in further analyses.

201 Nanopore community profiles by clinical group

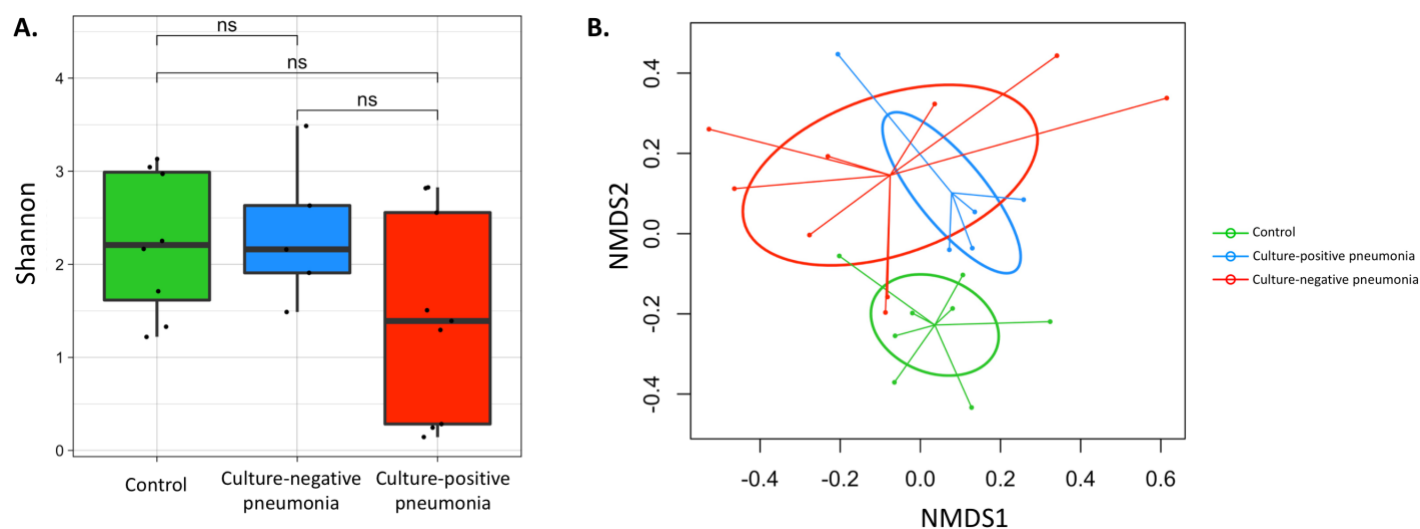
202 By Nanopore sequencing, culture-positive samples had a trend for lower alpha diversity (Shannon index)
203 compared to culture-negative samples (Fig 2A and Fig S3A) and demonstrated global compositional

204 dissimilarities compared to culture-negative and control samples (PERMANOVA p-value=0.038, $R^2=0.12$, Fig
205 2B and Fig S3B).

206

207 **Fig 2. Comparisons of lung microbial communities between samples of culture-positive pneumonia,**
208 **culture-negative pneumonia and controls based on Nanopore metagenomic sequencing.** (A) Compared
209 to samples from patients with culture-negative pneumonia, culture-positive samples had a trend for lower alpha
210 diversity of lung microbial communities by Shannon index. (B) By non-metric multidimensional scaling (NMDS)
211 plot of the Bray-Curtis dissimilarity index, there were significant differences in overall microbial community
212 compositions between three groups (PERMANOVA for Bray-Curtis dissimilarity index, p=0.0378, $R^2=0.120$).

213



214

215

216 Nanopore-based pathogen identification

217 Culture-positive pneumonia

218

219 We first examined Nanopore results in the culture-positive cases in which microbiologic confirmation of
220 the causative pathogen allowed for a targeted interrogation of the sequencing output for the corresponding
221 microbial species. We examined different thresholds of sequencing output (i.e. absolute number of reads for the
222 dominant pathogen vs. relative or ranked abundance thresholds for pathogens) to maximize sensitivity of
Nanopore results for detecting the culture-identified pathogen(s). By focusing on the three most abundant

223 species (bacterial or fungal) by Nanopore sequencing, we were able to identify all culture-confirmed pathogens
224 with high relative abundances.

225 In eight culture-positive bacterial pneumonias, Nanopore profiles showed high abundance of the same
226 bacterial species isolated in cultures (Fig 3A). These highly abundant causative bacterial pathogens had on
227 average 90 times higher relative abundance compared to the species ranked second in abundance in each
228 community (Fig 3B). Nanopore sequencing also revealed high abundance of additional potential bacterial
229 pathogens in 2/8 of samples that were not detected by cultures (*E. coli* in subject 1 and *H. influenzae* in subject
230 8), suggesting the presence of a polymicrobial infection despite the isolation of a single pathogenic bacterial
231 species on standard cultures (Fig 3A).

232 These eight culture-positive cases with clinical antibiograms allowed for examination of the potential
233 predictive utility of antibiotic resistance gene detection with metagenomic sequencing (Table 2). In the single
234 case of methicillin-resistant *Staphylococcus aureus* (MRSA, case 2, Fig S3), Nanopore detected 4 reads aligned
235 to the responsible *mecA* gene, whereas in the three cases of methicillin-sensitive *S. aureus* (MSSA, cases 3, 5
236 and 6), no *mecA* gene reads were detected. Similarly, in the two cases of *Stenotrophomonas maltophilia* and
237 *E.coli*, Nanopore detected genes that explained the observed phenotypic antimicrobial resistance profile (Table
238 2).

239 We also tested the performance of Nanopore sequencing in one case with probable invasive fungal
240 infection. Subject 9 was a lung transplant recipient who had been receiving antifungal therapy for a positive
241 sputum culture for *Aspergillus fumigatus*. Clinical decompensation with acute respiratory failure raised concern
242 for bacterial co-infection and initiation of intensive broad-spectrum antibiotics. BALF culture grew again
243 *Aspergillus fumigatus*, which was the dominant pathogen detected by Nanopore, without any other sequencing
244 evidence of bacterial infection (Fig 3C). Thus, in this case with probable invasive fungal infection, Nanopore
245 sequencing only identified a confirmed fungal pathogen and ruled out the presence of bacterial pneumonia.

246

247 Culture-negative pneumonia

248 Nanopore sequencing provided diverse representations of microbial communities in cases of clinical
249 suspicion of pneumonia with negative cultures, and thus interpretation needs to be individualized for each case
250 (Fig S3).

251 In case 10 with an initial diagnosis of aspiration pneumonia caused by *S. aureus* and *Klebsiella oxytoca*
252 identified by BALF culture on day 2 post-intubation, clinical deterioration on appropriate antibiotic therapy and
253 new radiographic infiltrates by day 5 raised concern for VAP. However, repeat BAL culture on day 5 was
254 negative. Nanopore detected high abundance of both *S. aureus* and *Klebsiella oxytoca* on day 5 sample,
255 revealing that the culture-negative community consisted of abundant previously identified pathogens, which were
256 likely not viable at the time of the day 5 BALF acquisition due to ongoing antibiotic therapy. Moreover, low
257 procalcitonin level at the time of the day 5 sample (94 pg/ μ l) and absence of new pathogens by sequencing
258 made diagnosis of new VAP unlikely.

259 In another case of a lung transplant recipient (subject 11) with diffuse bilateral consolidations and
260 persistent clinical septic picture of undefined etiology, BALF culture was only positive for yeast and the patient
261 was empirically treated with broad-spectrum antibiotics. Eventually, the patient was proven to be fungemic with
262 delayed growth of *Candida glabrata* on initial blood cultures prompting addition of antifungal therapy. Nanopore
263 sequencing on ETA sample from day 1 post-intubation showed high abundance of *Candida glabrata* and
264 *Candida dubliniensis* with very low abundance of bacterial reads, confirming the absence of bacterial pneumonia
265 and demonstrating fungal colonization of the allograft. Of note, in two other culture-negative cases, Nanopore
266 also detected high abundance of *Candida albicans* (Fig 3D), whereas in the last case, both Nanopore and 16S
267 sequencing identified abundant oral bacteria with no typical respiratory pathogens.

268

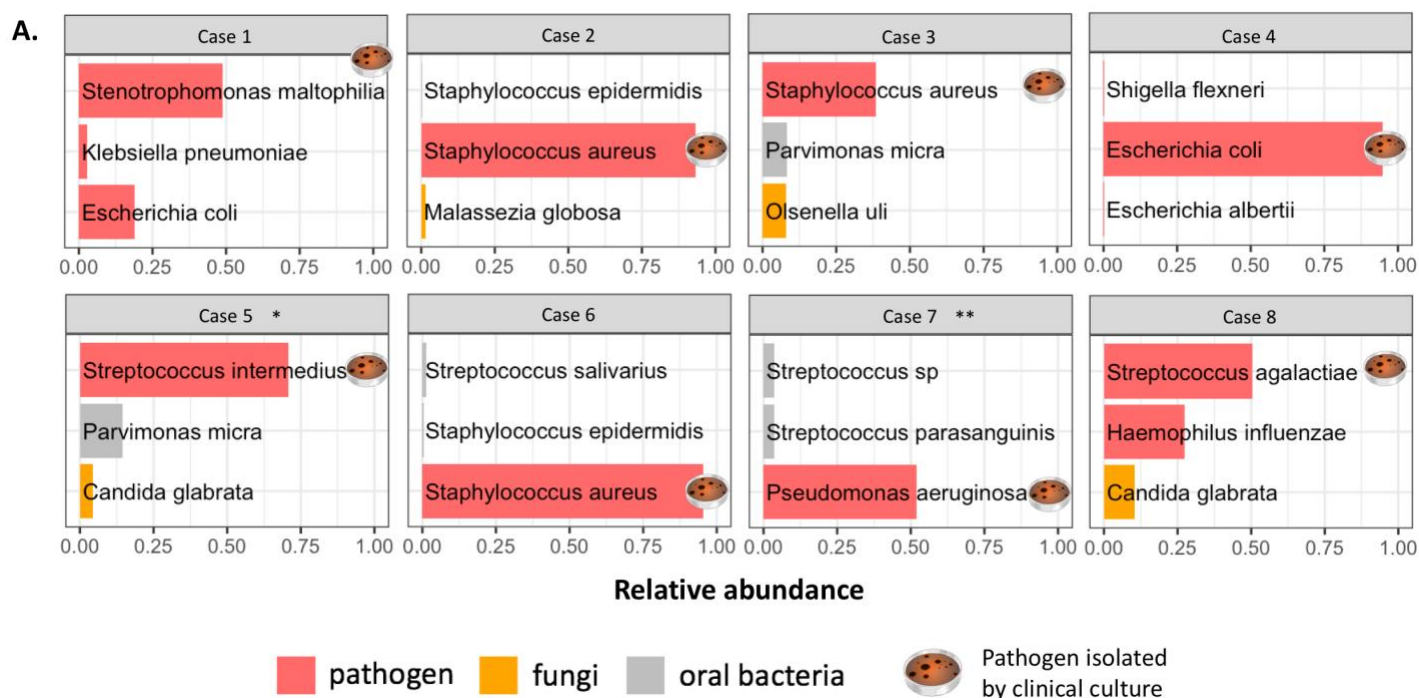
269 Controls

270 Eight control subjects did not meet clinical diagnostic criteria for pneumonia on retrospective examination
271 of their clinical course. Despite not meeting diagnostic criteria, 5/8 cases empiric antibiotics were prescribed
272 empiric antibiotics early in their course for initial consideration of possible pneumonia. Five samples were
273 dominated by common oral bacteria, such as *Rothia* and non-*pneumoniae* *Streptococcus* species (31,32).
274 However, in the other three samples, Nanopore and 16S sequencing detected potential respiratory pathogens
275 (e.g. *S.aureus*, *H. influenzae* or *S. pneumoniae*) that were likely airway colonizers not causing clinical infection,
276 notion supported by clinical improvement despite early discontinuation of antibiotics and/or low procalcitonin
277 levels (Fig S3). No significant fungal DNA presence was detected by Nanopore sequencing in the control group
278 (Fig 3E).

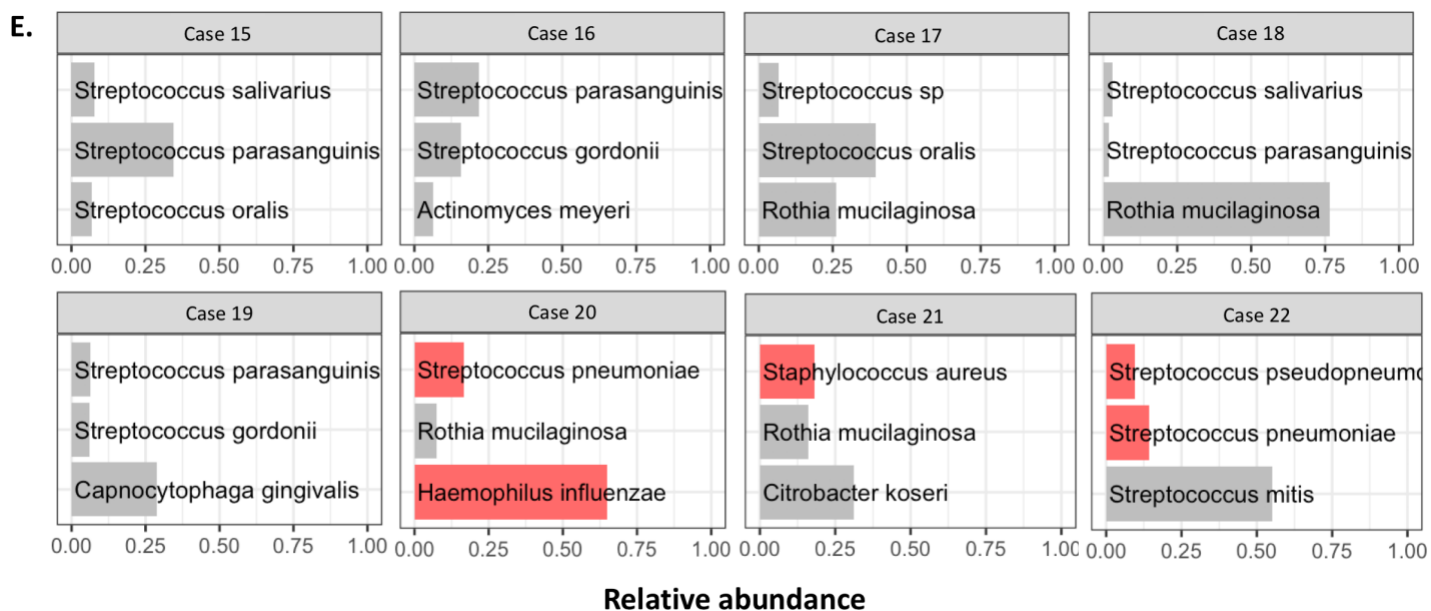
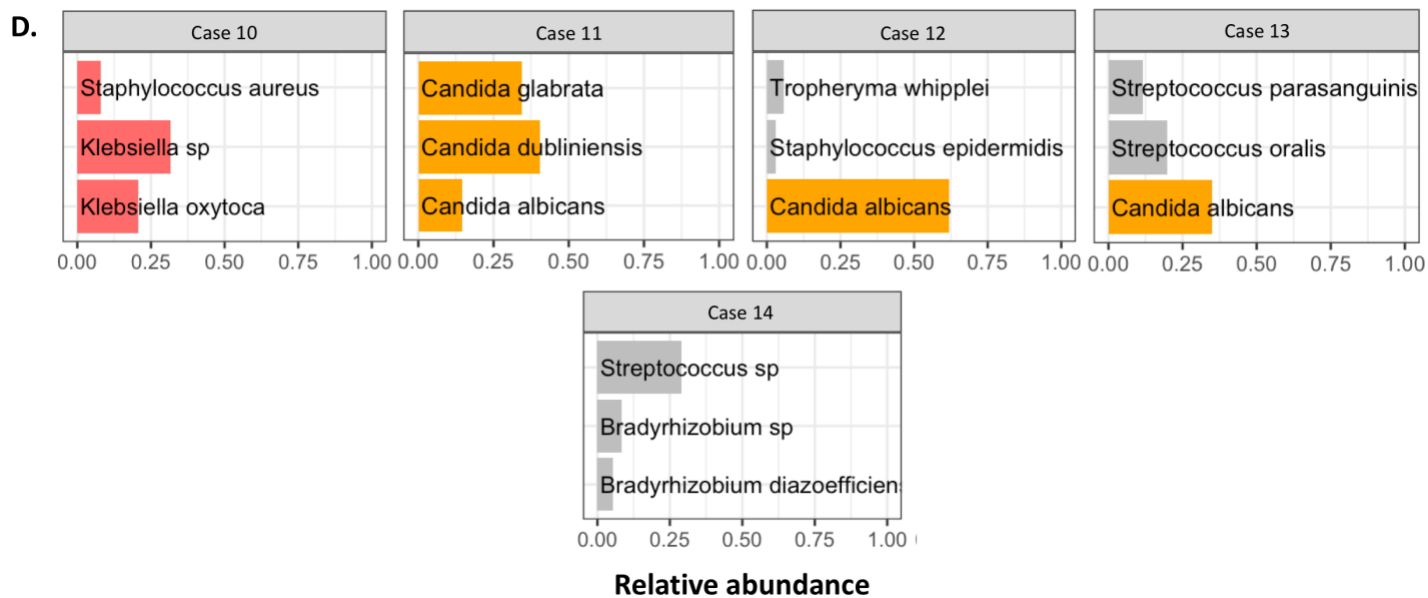
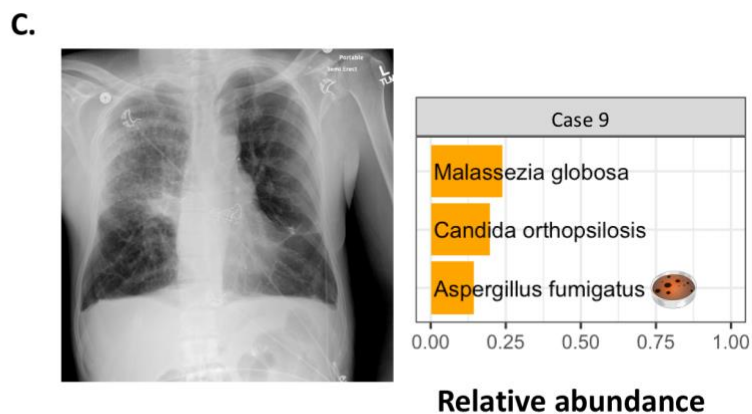
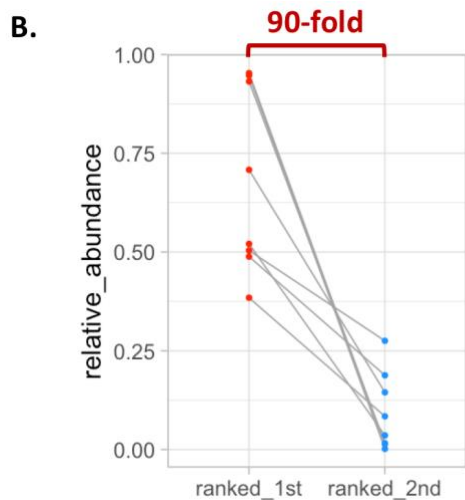
279 **Fig 3. Comparisons of microbes detected by Nanopore metagenomic sequencing and clinical culture.**

280 Each small plot represents an endotracheal aspirate; Each bar represents a microbe; the X-axis represents the
 281 relative abundance of microbes by Nanopore. Petri dish represents pathogen isolated by clinical culture. The
 282 three most abundant taxa detected by Nanopore sequencing were included. (A) In 8 samples with culture-
 283 positive bacterial pneumonia, Nanopore signals were dominated by pathogens isolated by culture. (B) In 8
 284 samples with culture-positive bacterial pneumonia, the relative abundance of culture-positive pathogens was 90-
 285 times higher than that of the second-ranked taxa detected by Nanopore. (C) In 1 sample with probable invasive
 286 fungal infection, chest radiograph supported a clinical diagnosis of pneumonia, *Aspergillus fumigatus* was
 287 isolated by culture, and Nanopore revealed the same fungal pathogen by sequencing. (D) In 5 culture-negative
 288 pneumonia samples, potential pathogens were found in one sample, and fungi were found in 3 samples with
 289 Nanopore. (E) Only typical oral bacteria were identified in 5/8 of control samples, but potential pathogens were
 290 detected in 3/8 of them. * compared to culture of pleural fluid; ** case of culture-positive tracheobronchitis and
 291 acute exacerbation of chronic obstructive pulmonary disease (no infiltrate on chest radiograph)

292



293



297 **Table 2: Comparison of antibiotic resistance phenotype detected by clinical culture and clinically**
 298 **relevant resistance genes detected by Nanopore in cases of bacterial pneumonia.** Genes conferring
 299 resistance phenotype are highlighted in bold.

Sample_ID	Pathogen by culture & Nanopore	Resistance phenotype by culture	Clinically relevant resistance genes [N of alignments]
Resistance identified			
Case 1	<i>Stenotrophomonas maltophilia</i>	R: ticarcillin/clavulanic acid R: ceftazidime	<i>blaTEM-4</i> [4] <i>blaTEM-112</i> [1] <i>blaTEM-157</i> [1] <i>blaACT-5</i> [1]
		I: levofloxacin	<i>oqxB</i> [1]
		Tetracycline not tested	<i>tetC</i> [6]
Case 2	<i>Staphylococcus aureus</i>	R: methicillin	<i>mecA</i> [4]
		R: erythromycin, clindamycin	<i>ermA</i> [10] <i>erm(33)</i> [1]
			<i>tet38</i> [11] <i>ant(4')-Ib</i> [9] <i>tetC</i> [3] <i>blaTEM-4</i> [1]
Case 3	<i>Staphylococcus aureus</i>	I: tetracycline	<i>tetK</i> [1] <i>tet38</i> [1] <i>tetQ</i> [1]
Case 4	<i>Escherichia Coli</i>	R: tetracycline	<i>tetX</i> [1]
		R: Trimethoprim-sulfamethoxazole	<i>sul1</i> [363] <i>dfrA</i> [127]
		R: ciprofloxacin, levofloxacin	<i>acrF</i> [315] [§] <i>parE</i> [304] [§] <i>mfd</i> [277] [§]
			<i>mphA</i> [215] <i>aadA5</i> [196] <i>vgaC</i> [110] <i>blaACT-5</i> [7] <i>blaACT-14</i> [3] <i>mefA</i> [1] <i>mel</i> [1]
No resistance identified			
Case 5	<i>Staphylococcus aureus</i>	S: all tested agents	none
Case 6	<i>Staphylococcus aureus</i>	S: all tested agents	<i>tet38</i> [2] <i>blaTEM4</i> [2]
Case 7	<i>Pseudomonas aeruginosa</i>	S: all tested agents	none
Resistance not tested			

Case 8	<i>Streptococcus agalactiae</i>	Not tested	<i>tetM</i> [60] <i>isaC</i> [55] <i>sul1</i> [2] <i>tetQ</i> [2] <i>mphA</i> [1] <i>aadA5</i> [1]
--------	---------------------------------	------------	---

300

301 Abbreviations: R: resistant; I: intermediate; S: Susceptible

302 § Genes conferring antibiotic resistance phenotype but not classified as clinical relevant genes by EPI2ME

303 antimicrobial resistance gene analyses [ARMA workflow]

304 Tested agents for case 5: Ampicillin/Sulbactam, Oxacillin, Imipenem, Gentamicin, Erthromycin, Tetracycline,

305 Vancomycin, Clindamycin, Linezolid, Rifampin, Trimethoprim-sulfamethoxazole, Synercid.

306 Tested agents for case 6: Ampicillin/Sulbactam, Oxacillin, Imipenem, Gentamicin, Erthromycin, Tetracycline.

307 Tested agents for case 7: Piperacillin/Tazobactam, Ticarcillin/Clavulanic acid, Cefepime, Ceftazidime,

308 Imipenem, Meropenem, Aztreonam, Gentamicin, Tobramycin, Amikacin, Ciprofloxacin, Levofloxacin.

309

310 Discussion

311 In this nested case-control study, we provide proof-of-concept evidence that untargeted, shotgun

312 metagenomic sequencing with the MinION device can provide clinically useful information for etiologic diagnosis

313 of pneumonia in mechanically-ventilated patients. We demonstrate feasibility of metagenomic sequencing

314 directly from clinical respiratory specimens by applying a saponin-based protocol for human DNA depletion prior

315 to sequencing. Our analyses demonstrated global microbial community structure and species-level

316 compositional differences associated with culture-positivity and clinical diagnosis of pneumonia. Nanopore

317 sequencing had striking concordance with cultures by detecting high abundance of the causative pathogenic

318 bacteria in culture-positive cases and refuting bacterial pneumonia diagnosis in selected culture-negative cases.

319 Nanopore metagenomic sequencing holds promise as a potential infection diagnostic tool due to its

320 comprehensive scope, high resolution and real-time data generation (6). However, contaminating human DNA

321 has been a rate-limiting step for clinical implementation of direct-from-sample sequencing in respiratory

322 specimens. By applying a recently validated protocol with saponin-based, human DNA depletion (13), we

323 demonstrate that this approach is feasible, reproducible and effective for maximizing the microbial signal in

324 clinical samples and providing reaching reliable diagnostic output. With further technical improvements, process
325 automation and cost reductions of metagenomic sequencing, this approach could be introduced in the clinical
326 microbiology laboratories as part of the diagnostic pipeline.

327 Nanopore sequencing showed high accuracy in pathogen identification in culture-positive pneumonia.
328 Obviating the need for *ex-vivo* growth for organisms, direct-from-sample sequencing can offer comprehensive
329 snapshots of the component microbes of the communities at the time of sample acquisition. Sequencing methods
330 are thus robust to specific growth condition requirements or the impaired viability of organisms due to antecedent
331 antimicrobials, factors that can lead to the 'great plate count anomaly' of culture-based methods, i.e. the absent
332 or limited growth of bacteria in cultures despite abundant visualized organisms on specimen gram-staining (33).
333 In exploratory analyses of the sequencing output, Nanopore also provided antibiotic resistance information by
334 detecting clinically relevant resistance genes that matched the phenotypic resistance on antibiograms (e.g. *mecA*
335 detection/absence in MRSA/MSSA cases, respectively). Overall, rapid metagenomic sequencing closely
336 matched the results of current, reference-standard diagnostic methods in our cohort, which typically take 2-3
337 days for allowing appropriate antibiotic adjustments to occur. Thus, with further external validation in additional
338 cohorts, nanopore metagenomics hold promise for rapidly accelerating the diagnosis of etiologic pathogens and
339 shortening the time to appropriate therapy selection in cases where diagnosis currently relies on cultures.

340 Invasive pulmonary fungal infections represent a major diagnostic challenge due to the poor sensitivity
341 and slow turnaround times of cultures and the need for invasive samples with histopathology for diagnostic
342 confirmation (34). In the single case of *Aspergillus fumigatus* infection, Nanopore confirmed the high abundance
343 of *Aspergillus fumigatus* in the community and ruled out concomitant bacterial pneumonia. Such results can
344 directly influence clinical practice, as unnecessary and potentially harmful antibiotics could be discontinued with
345 antimicrobial therapies focused on the causative fungal pathogen (35).

346 In cases of culture-negative pneumonia and in controls, the main compositional pattern consisted of
347 diverse communities with oral bacteria abundance (5), similar to clinical microbiology reports of normal
348 respiratory flora. In such cases, early de-escalation or discontinuation of antibiotics could be further supported
349 by sequencing results, if available in real-time. Nonetheless, in 3/5 airway controls, potential respiratory
350 pathogens were detected in high abundance by both Nanopore and 16S sequencing, in the absence of other
351 supportive evidence of pneumonia. Such cases highlight that the high sensitivity of sequencing for identifying

352 pathogenic organisms missed by cultures could accentuate the existing clinical challenge of distinguishing
353 colonizing vs. infecting organisms in the airways of mechanically-ventilated patients. In such scenarios, the
354 distinction between colonization and infection cannot be based solely on microbial DNA sequencing outputs, but
355 needs to be an integrative one, incorporating clinical, radiographic, and systemic or focal host-responses (36–
356 38). At the same time, knowledge of colonizing organisms in critically-ill patients could facilitate more targeted
357 choices for initial antibiotic regimens in the event of a secondary infection, such as VAP.

358 Our study is limited by the single center design and the available sample size. We did not perform
359 Nanopore sequencing and data analyses in real-time because of our retrospective study design, and our
360 objective of demonstrating proof-of-concept feasibility. Nonetheless, the method is implementable with short
361 turnaround times (~6-8hrs) (13). The results of antibiotic resistance gene sequencing should be interpreted
362 with caution, given that our analyses were exploratory, included a limited number of multidrug-resistant
363 bacteria that precluded a formal predictive modeling analysis. Thus, development of reliable predictive rules
364 for pneumonia diagnosis or resistance gene identification based on sequencing outputs will require
365 prospective examination of large cohorts of patient samples. Finally, the human DNA depletion method we
366 applied is not yet optimized for viral DNA detection (16). Nonetheless, most clinically relevant respiratory
367 viruses are RNA organisms, which are not within the scope of current DNA-based metagenomics, but can be
368 detected with available PCR-based panels. With evolving rapid metagenomic platforms that can also
369 sequence RNA molecules (39), the profiling of respiratory viruses as well as host transcriptomic responses
370 (37) will enable more comprehensive representations of the altered respiratory ecosystem in pneumonia.

371 In conclusion, our study demonstrates the technical feasibility and clinical validity of direct-from-specimen
372 metagenomics with a rapid protocol for human DNA depletion protocol and sequencing with the MiNION device.
373 Metagenomic approaches hold promise for the development of rapid and comprehensive diagnostic tools for
374 severe pneumonia that could transform the existing diagnostic paradigm. With real-time data generation and
375 turnaround times of 6-8hrs from sample acquisition to result, rapid metagenomics could conceivably allow for
376 targeted adjustment of initial empiric antibiotic regimens even before their second dose is due, and thus allow
377 for personalized antimicrobial prescriptions and antibiotic stewardship gains. Our results provide strong rationale
378 for a prospective, large-scale study with real-time application of metagenomics in order to measure the direct
379 impact on antibiotic guidance and clinical outcomes.

380

381 Data Sharing Statements: All de-identified sequencing data were submitted to Sequence Read Archive (SRA)
382 database, accession numbers 12268279 – 12268349. All de-identified datasets for this study are provided in
383 <https://github.com/MicrobiomeALIR>.

384

385 **Acknowledgement:** We would like to thank all members of the research team of the Acute Lung Injury Registry
386 (ALIR) and Biospecimen Repository at the University of Pittsburgh, the medical and nursing staff in the Medical
387 Intensive Care Unit at the University of Pittsburgh Medical Center, and all patients and their families for
388 participating in this research project.

389

390

391 **References:**

392

- 393 1. Xu J, Murphy SL, Kochanek KD, Bastian B, Arias E. Deaths: final data for 2016. *Natl Vital Stat Rep*. 2018
394 Jul;67(5):1–76.
- 395 2. Vaughn VM, Flanders SA, Snyder A, Conlon A, Rogers MAM, Malani AN, et al. Excess antibiotic
396 treatment duration and adverse events in patients hospitalized with pneumonia: A multihospital cohort
397 study. *Ann Intern Med*. 2019 Jul 9;
- 398 3. Jain S, Self WH, Wunderink RG, Fakhran S, Balk R, Bramley AM, et al. Community-Acquired Pneumonia
399 Requiring Hospitalization among U.S. Adults. *N Engl J Med*. 2015 Jul 30;373(5):415–27.
- 400 4. Zumla A, Al-Tawfiq JA, Enne VI, Kidd M, Drosten C, Breuer J, et al. Rapid point of care diagnostic tests
401 for viral and bacterial respiratory tract infections--needs, advances, and future prospects. *Lancet Infect
402 Dis*. 2014 Nov;14(11):1123–35.
- 403 5. Kitsios GD, Fitch A, Manatakis DV, Rapport SF, Li K, Qin S, et al. Respiratory microbiome profiling for
404 etiologic diagnosis of pneumonia in mechanically ventilated patients. *Front Microbiol*. 2018 Jul 10;9:1413.
- 405 6. Dunlap DG, Marshall CW, Fitch A, Rapport SF, Cooper VS, McVerry BJ, et al. Improved detection of
406 culprit pathogens by bacterial DNA sequencing affects antibiotic management decisions in severe
407 pneumonia. *Am J Case Rep*. 2018 Nov 26;19:1405–9.
- 408 7. Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, et al. Clinical metagenomic sequencing
409 for diagnosis of meningitis and encephalitis. *N Engl J Med*. 2019 Jun 13;380(24):2327–40.
- 410 8. Dickson RP, Erb-Downward JR, Prescott HC, Martinez FJ, Curtis JL, Lama VN, et al. Analysis of culture-
411 dependent versus culture-independent techniques for identification of bacteria in clinically obtained
412 bronchoalveolar lavage fluid. *J Clin Microbiol*. 2014 Oct;52(10):3605–13.
- 413 9. Kalantar KL, Moazed F, Christenson SC, Wilson J, Deiss T, Belzer A, et al. Metagenomic comparison of
414 tracheal aspirate and mini-bronchial alveolar lavage for assessment of respiratory microbiota. *Am J
415 Physiol Lung Cell Mol Physiol*. 2019 Mar 1;316(3):L578–84.
- 416 10. Kelly BJ, Imai I, Bittinger K, Laughlin A, Fuchs BD, Bushman FD, et al. Composition and dynamics of the
417 respiratory tract microbiome in intubated patients. *Microbiome*. 2016 Feb 11;4:7.
- 418 11. Poretsky R, Rodriguez-R LM, Luo C, Tsementzi D, Konstantinidis KT. Strengths and limitations of 16S
419 rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLoS One*. 2014
420 Apr 8;9(4):e93827.
- 421 12. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome
422 sequencing for Ebola surveillance. *Nature*. 2016 Feb 11;530(7589):228–32.
- 423 13. Faria NR, Sabino EC, Nunes MRT, Alcantara LCJ, Loman NJ, Pybus OG. Mobile real-time surveillance
424 of Zika virus in Brazil. *Genome Med*. 2016 Sep 29;8(1):97.
- 425 14. Schmidt K, Mwaigwisya S, Crossman LC, Doumith M, Munroe D, Pires C, et al. Identification of bacterial
426 pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic
427 sequencing. *J Antimicrob Chemother*. 2017;72(1):104–14.
- 428 15. Pendleton KM, Erb-Downward JR, Bao Y, Branton WR, Falkowski NR, Newton DW, et al. Rapid
429 Pathogen Identification in Bacterial Pneumonia Using Real-Time Metagenomics. *Am J Respir Crit Care
430 Med*. 2017 Dec 15;196(12):1610–2.
- 431 16. Charalampous T, Kay GL, Richardson H, Aydin A, Baldan R, Jeanes C, et al. Nanopore metagenomics

- 432 enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat Biotechnol.* 2019 Jun 24;
- 433 17. Kitsios GD, Yang L, Manatakis DV, Nouraie M, Evankovich J, Bain W, et al. ARDS subphenotypes
434 beyond ARDS: prognostic enrichment in mechanically-ventilated patients with or at risk for ARDS. *Crit*
435 *Care Med.*
- 436 18. Gong MN, Thompson BT, Williams P, Pothier L, Boyce PD, Christiani DC. Clinical predictors of and
437 mortality in acute respiratory distress syndrome: potential role of red cell transfusion. *Crit Care Med.*
438 2005 Jun;33(6):1191–8.
- 439 19. Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing.
440 25th informational supplement, M100-S25. Wayne, PA. 2018;
- 441 20. Zilberberg MD, Shorr AF. Ventilator-associated pneumonia: the clinical pulmonary infection score as a
442 surrogate for diagnostics and outcome. *Clin Infect Dis.* 2010 Aug 1;51 Suppl 1:S131-5.
- 443 21. Cookson WOCM, Cox MJ, Moffatt MF. New opportunities for managing acute and chronic lung
444 infections. *Nat Rev Microbiol.* 2018 Feb;16(2):111–20.
- 445 22. Liu CM, Aziz M, Kachur S, Hsueh P-R, Huang Y-T, Keim P, et al. BactQuant: an enhanced broad-
446 coverage bacterial quantitative real-time PCR assay. *BMC Microbiol.* 2012 Apr 17;12:56.
- 447 23. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput
448 microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 2012 Aug;6(8):1621–
449 4.
- 450 24. Morgulis A, Gertz EM, Schäffer AA, Agarwala R. A fast and symmetric DUST implementation to mask
451 low-complexity DNA sequences. *J Comput Biol.* 2006 Jun;13(5):1028–40.
- 452 25. Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A, et al. Manipulation of
453 FASTQ data with Galaxy. *Bioinformatics.* 2010 Jul 15;26(14):1783–5.
- 454 26. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* 2011
455 May 2;17(1):10.
- 456 27. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-
457 source, platform-independent, community-supported software for describing and comparing microbial
458 communities. *Appl Environ Microbiol.* 2009 Dec;75(23):7537–41.
- 459 28. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA
460 sequences into the new bacterial taxonomy. *Appl Environ Microbiol.* 2007 Aug;73(16):5261–7.
- 461 29. Fair K, Dunlap DG, Fitch A, Morris A, McVerry BJ, Kitsios G. Rectal swabs in critically-ill patients provide
462 discordant representations of the gut microbiome compared to stool samples: a brief methodologic
463 report. *BioRxiv.* 2019 Feb 5;
- 464 30. Dixon P. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science.* 2003
465 Dec;14(6):927–30.
- 466 31. Segal LN, Clemente JC, Tsay J-CJ, Koralov SB, Keller BC, Wu BG, et al. Enrichment of the lung
467 microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat Microbiol.* 2016
468 Apr 4;1:16031.
- 469 32. Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Falkowski NR, Huffnagle GB, et al.
470 Bacterial topography of the healthy human lower respiratory tract. *MBio.* 2017 Feb 14;8(1).
- 471 33. Staley JT, Konopka A. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic
472 and terrestrial habitats. *Annu Rev Microbiol.* 1985;39:321–46.

- 473 34. Lease ED, Alexander BD. Fungal diagnostics in pneumonia. *Semin Respir Crit Care Med*. 2011 Dec
474 13;32(6):663–72.
- 475 35. Sidransky H, Pearl MA. Pulmonary Fungus Infections Associated with Steroid and Antibiotic Therapy. *Dis*
476 *Chest*. 1961 Jun;39(6):630–42.
- 477 36. Kitsios GD. Translating Lung Microbiome Profiles into the Next-Generation Diagnostic Gold Standard for
478 Pneumonia: a Clinical Investigator’s Perspective. *mSystems*. 2018 Apr;3(2).
- 479 37. Langelier C, Kalantar KL, Moazed F, Wilson MR, Crawford ED, Deiss T, et al. Integrating host response
480 and unbiased microbe detection for lower respiratory tract infection diagnosis in critically ill adults. *Proc*
481 *Natl Acad Sci USA*. 2018 Dec 26;115(52):E12353–62.
- 482 38. Walter JM, Ren Z, Yacoub T, Reyfman PA, Shah RD, Abdala-Valencia H, et al. Multidimensional
483 Assessment of the Host Response in Mechanically Ventilated Patients with Suspected Pneumonia. *Am J*
484 *Respir Crit Care Med*. 2019 May 15;199(10):1225–37.
- 485 39. Depledge DP, Srinivas KP, Sadaoka T, Bready D, Mori Y, Placantonakis DG, et al. Direct RNA
486 sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen. *Nat*
487 *Commun*. 2019 Feb 14;10(1):754.

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505