

1 **Metagenomic sequencing to detect respiratory viruses in persons under**
2 **investigation for COVID-19**

3
4 Ahmed Babiker, MBBS^{1,2}; Heath L. Bradley, MSc²; Victoria D. Stittleburg, BS¹; Autum
5 Key, MSc²; Colleen S. Kraft MD, MSc^{1,2}; Jesse J. Waggoner, MD¹, Anne Piantadosi,
6 MD, PhD^{1,2}

7 1- Division of Infectious Diseases, Department of Medicine, Emory University
8 School of Medicine Atlanta, Georgia, USA

9 2- Department of Pathology and Laboratory Medicine, Emory University School of
10 Medicine Atlanta, Georgia, USA

11

12 **Running title:** Metagenomic sequencing in COVID-19 PUI

13 **Key words:** SARS-CoV-2; COVID-19; metagenomic sequencing; diagnostics;
14 respiratory viruses

15

16 **Corresponding Author:**

17 Anne Piantadosi, MD, PhD

18 Department of Pathology and Laboratory Medicine, Emory University School of
19 Medicine Atlanta, Georgia, USA

20 Email: anne.piantadosi@emory.edu

21

22 **Alternative corresponding author:**

23 Ahmed Babiker

24 Division of Infectious Diseases, Department of Medicine, Emory University School of

25 Medicine Atlanta, Georgia, USA

26 Email: ahmed.babiker@emory.edu

27

28

29 **Abstract**

30 We used metagenomic next-generation sequencing (mNGS) to assess the frequencies
31 of alternative viral infections in SARS-CoV-2 RT-PCR negative persons under
32 investigations (PUIs) (n=30) and viral co-infections in SARS-CoV-2 RT-PCR positive
33 PUIs (n=45). mNGS identified both co-infections and alternative viral infections that
34 were not detected by routine clinical workup.

35 **Introduction**

36 The coronavirus disease 2019 (COVID-19) pandemic caused by severe acute
37 respiratory syndrome coronavirus 2 (SARS-CoV-2) has precipitated a massive global
38 health and economic crisis, claiming over a quarter million lives since its emergence in
39 Wuhan, China in December, 2019 (1). Broad testing for respiratory viruses among
40 persons under investigation for SARS-CoV-2 is performed inconsistently, limiting our
41 understanding of alternative viral infections and co-infections in these patients. We
42 evaluated the performance of metagenomic next-generation sequencing (mNGS) to
43 detect SARS-CoV-2, co-infections, and alternative viral infections in persons under
44 investigation (PUIs) for SARS-CoV-2 during the first two months of the pandemic in the
45 state of Georgia.

47 **Materials and Methods**

48 A convenience sample set was selected from 75 PUIs who were tested for
49 SARS-CoV-2 in the Emory Healthcare system between February 26th and April 23rd
50 2020 (spanning the first detection of SARS-CoV-2 infection in Georgia on March 2nd and
51 in the Emory Healthcare system on March 3rd). Clinical data including laboratory results
52 were extracted by chart review. Flu/RSV PCR (Cepheid, Sunnyvale, CA), and multiplex
53 respiratory panels including the eSensor® Respiratory Viral Panel (GenMark
54 Diagnostics, Inc., Carlsbad, CA), BioFire® FilmArray® Biofire Respiratory pathogen and
55 BioFire® FilmArray® Pneumonia panel (BioFire Diagnostics, LLC, Salt Lake City, UT)
56 had been performed at the discretion of treating physicians. Excess nasopharyngeal
57 swab samples were retrieved within 72 hours of collection and underwent RNA

58 extraction and SARS-CoV-2 testing by triplex RT-PCR (2). Samples underwent
59 unbiased mNGS as previously described (3, 4). Briefly, this included DNase treatment,
60 random primer cDNA synthesis, Nextera XT tagmentation, and Illumina sequencing to a
61 median depth of 35 million reads. Reads underwent taxonomic classification by
62 KrakenUniq, as implemented in viral-ngs (5). Reads that were taxonomically assigned
63 to a human respiratory virus were extracted, aligned to a reference sequence in
64 Geneious Prime V2020.1.1, and manually inspected. Detection of a virus was reported
65 as confirmed (**Supplementary Tables 1 and 2**) if nonoverlapping reads from ≥ 3 distinct
66 genomic regions were identified (6). Reference-based viral genome assembly was
67 performed using viral-ngs. The following references were used for alignment and
68 assembly: NC_001803.1 (Respiratory Syncytial Virus (RSV)), NC_0391991.1 (Human
69 Metapneumovirus), NC_006577.2 (Human Coronavirus (HCoV) HKU1), NC_005831
70 (HCoV NL63), NC_006213.1 (HCoV OC43), and NC_026431.1-NC_026438.1
71 (Influenza H1N1). In some cases, reads were assigned to a virus by KrakenUniq but
72 failed to align to the reference; these reads underwent classification by BLASTn. All
73 reads (cleaned of human reads) will be submitted to the Sequence Read Archive and
74 complete genomes will be submitted to GenBank.

75

76 **Results**

77 75 PUIs were included, 30 of whom were negative and 45 of whom were positive
78 for SARS-CoV-2 by RT-PCR. Among the 30 SARS-CoV-2 negative PUIs, 93% (28/30)
79 underwent clinical testing for alternative infections, 73% (22/30) by Flu/RSV PCR testing
80 and 80% (24/36) by multiplex respiratory viral panel testing. Eight of the 30 SARS-CoV-

81 2 negative PUIs (27%) tested positive for another respiratory virus by routine clinical
82 testing. In all cases, mNGS identified the same virus (**Table 1**). Importantly, mNGS also
83 identified four viruses that were not detected by routine clinical testing: two
84 (Metapneumovirus and HCoV NL63) that were not tested for and two (RSV and
85 influenza, in the same patient sample) that were missed despite testing with a multiplex
86 respiratory viral panel (FilmArray Biofire Respiratory pathogen). In addition to these
87 confirmed findings, mNGS identified one read of HCoV NL63 in an otherwise negative
88 sample, which was below the limit of detection for this study, as well as reads initially
89 classified as alphacoronaviruses in 3 samples, which were ultimately determined to be
90 non-viral contaminants by BLASTn. No SARS-CoV-2 was detected by mNGS among
91 RT-PCR negative PUIs (**Supplementary Table 1**).

92 Among the 45 SARS-CoV-2 RT-PCR positive PUIs, 31 (69%) underwent clinical
93 testing for co-infections, and none were detected (**Table 1**). mNGS identified one viral
94 co-infection (RSV) that was not tested for; this was confirmed with a second
95 independent sequencing library. SARS-CoV-2 was detected by mNGS in all 45 RT-PCR
96 positive PUIs (**Supplementary Table 2**).

97

98 **Discussion**

99 Overall, using unbiased mNGS, we were able to readily identify SARS-CoV-2,
100 co-infections, and alternative infections in PUIs, including viruses not identified by
101 routine clinical testing. Reassuringly, we did not detect SARS-CoV-2 by mNGS in
102 samples that were negative by RT-PCR, indicating both that the RT-PCR assay was
103 sensitive, and that there were no divergent viruses missed due to primer mismatch (7).

104 Interestingly, many PUIs had no infection identified on either routine testing or mNGS,
105 which may reflect inadequate sampling, rapid clearance of a potential viral infection, or
106 a non-viral process. Limitations of our study include the testing of banked samples
107 which may degrade over time with freeze-thaw cycles, and the use of clinical tests (such
108 SARS-CoV-2 RT-PCR) as an imperfect gold standard. Nevertheless, our results offer
109 several valuable insights regarding molecular testing for respiratory viruses among
110 SARS-CoV-2 PUIs.

111 First, we found a low rate of viral co-infection in patients with SARS-CoV-2, with
112 only one co-infection (with RSV) among 45 patients (2%). Our results are concordant
113 with a systematic review of 16 studies and 1,014 patients, which found a 3% rate of viral
114 co-infection, with RSV and Influenza A being the most common (8). Thus, viral co-
115 infections may be less common than initially reported (9); nevertheless, it is imperative
116 to continue to monitor for them in the setting of the upcoming respiratory virus season
117 and ongoing waves of COVID-19.

118 Second, our results illustrate key considerations about the potential use of mNGS
119 as a diagnostic tool. Notably, mNGS detected all viruses that were identified by routine
120 clinical workup, as well as five viruses that were not identified by clinical workup. Three
121 of these five had not been tested for clinically, underscoring the importance of routine
122 broad-spectrum molecular testing for respiratory viruses among PUIs. However, two of
123 the viruses were tested for and not identified by multiplex PCR, suggesting that mNGS
124 may offer improved sensitivity. Additional advantages of mNGS have been reviewed
125 elsewhere (10) and include the opportunities to characterize viral genomics, the
126 microbiome, and the host transcriptome.

127 However, current challenges with implementing mNGS into routine clinical
128 microbiology workflows include its relatively high cost and low throughput, as well as the
129 need for expertise in interpreting results. One barrier in particular is the lack of
130 established thresholds for identifying pathogens. Herein, we required detection of
131 nonoverlapping reads from ≥ 3 distinct genome regions based upon previously
132 described criteria (6). Recently, Peddu *et al.* used a cutoff based on reads per million
133 RPM (< 10) (11); in our study, the use of this cutoff would have excluded viruses that
134 were found by both mNGS and routine clinical testing by PCR, which we regard as true
135 positives. Thus, there is ongoing need to harmonize practices around clinical mNGS
136 testing.

137 In conclusion, our results suggest that current broad-spectrum molecular testing
138 algorithms identify most respiratory viral infections among SARS-CoV-2 PUIs, when
139 available and implemented consistently. In addition, these results illustrate the potential
140 of mNGS to streamline and expand clinical testing for respiratory viruses, which may
141 augment strategies to surveil for unexpected viral co-infections or the emergence of
142 divergent strains during periods of high transmission.

143

144 **Funding**

145 This study was supported by the Pediatric Research Alliance Center for Childhood
146 Infections and Vaccines and Children's Healthcare of Atlanta. The Yerkes NHP
147 Genomics Core is supported in part by NIH P51 OD011132 and sequencing data was
148 acquired on an Illumina NovaSeq6000 funded by NIH S10 OD026799. Salary support

149 was provided by the Doris Duke Charitable Foundation, Clinical Scientist Development
150 Award 2019089 (JJW) and NIH K08 AI139348 (AP).

151

152 **Acknowledgments**

153 We would like to acknowledge our laboratory colleagues at the Emory University
154 Healthcare Microbiology and Molecular laboratories who have worked tirelessly to
155 provide necessary care to our patients during this time. We would also like to thank the
156 staff of the Emory Integrated Genomics Core, Yerkes NHP Genomics Core and Jessica
157 Ingersoll of the Emory Clinical Virology Research Laboratory for sequencing support,
158 and Daniel J. Park for valuable advice on the use of viral-ngs.

159

160 **Table 1: Molecular and Metagenomic Testing of Persons Under Investigation**
 161

| | SARS-CoV-2 RT-PCR negative PUIs (N=30) | SARS-CoV-2 RT-PCR positive PUIs (N=45) |
|---|--|---|
| PUIs with molecular testing for non-SARS-CoV-2 respiratory viruses | 28 (93%) | 31 (69%) |
| Flu/RSV PCR | 22 (73%) | 28 (62%) |
| Multiplex Respiratory Pathogen PCR† | 24 (80%) | 21 (47%) |
| PUIs with alternative and co-infecting viruses detected by both routine methods and mNGS | 8 (29%) | 0 (0%) |
| Specific viruses detected by both routine methods and mNGS | Influenza A (N=3), Human metapneumovirus (N=2), Human coronavirus OC43 (N=2), Human coronavirus HKU1(N=1) | n/a |
| PUIs with additional viruses detected by mNGS alone | 3 (14%) | 1 (2%) |
| Specific viruses detected by mNGS alone | Respiratory syncytial virus (N=1), Human metapneumovirus (N=1), Human coronavirus NL63 (N=1), Influenza A (N=1) | RSV (N=1) |
| PUIs with SARS-CoV-2 detected by mNGS | 0 (0%) | 45 (100%) |

162 N (%) unless otherwise stated

163
 164 Abbreviations: mNGS: metagenomic next-generation sequencing, PCR: polymerase chain reaction, RT-PCR: reverse-transcriptase
 165 PCR, RSV: respiratory syncytial virus, SARS-CoV-2: severe acute respiratory syndrome-coronavirus-2

166
 167 † Multiplex respiratory panels included eSensor® Respiratory Viral Panel (GenMark Diagnostics, Inc., Carlsbad, CA) and BioFire®
 168 FilmArray® Respiratory and Pneumonia panels (BioFire Diagnostics, LLC, Salt Lake City, UT)

169

170

171

172

173 **References:**

174

- 175 1. Dong E, Du H, Gardner L. 2020. An interactive web-based dashboard to track
176 COVID-19 in real time. *Lancet Infect Dis* doi:10.1016/s1473-3099(20)30120-1.
- 177 2. Waggoner JJ, Stittleburg V, Pond R, Saklawi Y, Sahoo MK, Babiker A, Hussaini
178 L, Kraft CS, Pinsky BA, Anderson EJ, Rouphael N. 2020. Triplex Real-Time RT-
179 PCR for Severe Acute Respiratory Syndrome Coronavirus 2. *Emerging infectious*
180 *diseases* 26.
- 181 3. Piantadosi A, Freije CA, Gosmann C, Ye S, Park D, Schaffner SF, Tully DC,
182 Allen TM, Dong KL, Sabeti PC, Kwon DS. 2019. Metagenomic Sequencing of
183 HIV-1 in the Blood and Female Genital Tract Reveals Little Quasispecies
184 Diversity during Acute Infection. *Journal of virology* 93:e00804-18.
- 185 4. Piantadosi A, Mukerji SS, Chitneni P, Cho TA, Cosimi LA, Hung DT, Goldberg
186 MB, Sabeti PC, Kuritzkes DR, Grad YH. 2017. Metagenomic Sequencing of an
187 Echovirus 30 Genome From Cerebrospinal Fluid of a Patient With Aseptic
188 Meningitis and Orchitis. *Open forum infectious diseases* 4:ofx138-ofx138.
- 189 5. Park DJ, Dudas G, Wohl S, Goba A, Whitmer SLM, Andersen KG, Sealfon RS,
190 Ladner JT, Kugelman JR, Matranga CB, Winnicki SM, Qu J, Gire SK, Gladden-
191 Young A, Jalloh S, Nosamiefan D, Yozwiak NL, Moses LM, Jiang P-P, Lin AE,
192 Schaffner SF, Bird B, Towner J, Mamoh M, Gbakie M, Kanneh L, Kargbo D,
193 Massally JLB, Kamara FK, Konuwa E, Sellu J, Jalloh AA, Mustapha I, Foday M,
194 Yillah M, Erickson BR, Sealy T, Blau D, Paddock C, Brault A, Amman B, Basile J,
195 Bearden S, Belser J, Bergeron E, Campbell S, Chakrabarti A, Dodd K, Flint M,

- 196 Gibbons A, et al. 2015. Ebola Virus Epidemiology, Transmission, and Evolution
197 during Seven Months in Sierra Leone. *Cell* 161:1516-1526.
- 198 6. Miller S, Naccache SN, Samayoa E, Messacar K, Arevalo S, Federman S, Stryke
199 D, Pham E, Fung B, Bolosky WJ, Ingebrigtsen D, Lorizio W, Paff SM, Leake JA,
200 Pesano R, DeBiasi R, Dominguez S, Chiu CY. 2019. Laboratory validation of a
201 clinical metagenomic sequencing assay for pathogen detection in cerebrospinal
202 fluid. *Genome Res* 29:831-842.
- 203 7. Osório NS, Correia-Neves M. 2020. Implication of SARS-CoV-2 evolution in the
204 sensitivity of RT-qPCR diagnostic assays. *Lancet Infect Dis* doi:10.1016/s1473-
205 3099(20)30435-7.
- 206 8. Lansbury L, Lim B, Baskaran V, Lim WS. 2020. Co-infections in people with
207 COVID-19: a systematic review and meta-analysis. *The Journal of infection*
208 81:266-275.
- 209 9. Kim D, Quinn J, Pinsky B, Shah NH, Brown I. 2020. Rates of Co-infection
210 Between SARS-CoV-2 and Other Respiratory Pathogens. *JAMA* 323:2085-2086.
- 211 10. Chiu CY, Miller SA. 2019. Clinical metagenomics. *Nat Rev Genet* 20:341-355.
- 212 11. Peddu V, Shean RC, Xie H, Shrestha L, Perchetti GA, Minot SS, Roychoudhury
213 P, Huang M-L, Nalla A, Reddy SB, Phung Q, Reinhardt A, Jerome KR, Greninger
214 AL. 2020. Metagenomic analysis reveals clinical SARS-CoV-2 infection and
215 bacterial or viral superinfection and colonization. *Clinical Chemistry*
216 doi:10.1093/clinchem/hvaa106.
- 217