1	Title: Respiratory virus detection and sequencing from negative SARS-CoV-2 rapid antigen tests
2	

3 Authors:

- 4 Emmanuela Jules^{1,*}, Charlie Decker^{1,*}, Brianna Jeanne Bixler PhD^{1,2,*}, Alaa Ahmed PhD^{1,3},
- 5 Zijing (Carol) Zhou¹, Itika Arora PhD^{1,2}, Henok Tafesse MD¹, Hannah Dakanay¹, Andrei
- 6 Bombin PhD^{1,2}, Ethan Wang¹, Jessica Ingersoll¹, Kathy Bifulco⁴, Jennifer K. Frediani PhD,
- 7 RD^{4,5}, Richard Parsons^{4,5}, Julie Sullivan^{5,6}, Morgan Greenleaf^{5,7}, Jesse J. Waggoner MD², Greg
- 8 S. Martin MD, MSc⁸, Wilbur A. Lam MD, PhD^{6,9,10}, Anne Piantadosi MD, PhD^{1,2,#}
- 9

10 Affiliations:

- ¹Department of Pathology and Laboratory Medicine; Emory University School of Medicine;
- 12 Atlanta, GA, 30322; USA
- 13 ²Division of Infectious Diseases; Department of Medicine, Emory University School of
- 14 Medicine; Atlanta, GA, 30322; USA
- ³Emory Integrated Genomics Core; Emory University School of Medicine; Atlanta, GA, 30322;
- 16 USA
- 17 ⁴Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, Georgia, USA
- ⁵The Atlanta Center for Microsystems-Engineered Point-of-Care Technologies, Atlanta, GA,
- 19 USA
- ⁶Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA
- 21 ⁷Georgia Clinical and Translational Science Alliance, Atlanta, GA
- ⁸Division of Pulmonary, Allergy, Critical Care Medicine and Sleep Medicine, Emory University,
- 23 Atlanta, GA 30303, USA

- ⁹Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology
- and Emory University, Atlanta, GA, USA.
- 26 ¹⁰Aflac Cancer and Blood Disorders Center of Children's Healthcare of Atlanta, Atlanta, GA,
- 27 USA
- 28 *These first authors contributed equally to this article
- 29
- **30** *#* **Corresponding author**:
- 31 Anne Piantadosi MD, PhD
- 32 Woodruff Memorial Research Building
- 33 101 Woodruff Circle
- 34 Atlanta GA 30322
- 35 404-712-9005
- 36 <u>anne.piantadosi@emory.edu</u>
- 37
- 38 Word count:
- 39 Abstract: 45
- 40 Text: 1,199
- 41
- 42 **Running title:** Metagenomics from negative COVID-19 antigen tests
- 43
- 44 Keywords: molecular epidemiology, respiratory tract infections, public health surveillance,
- 45 SARS-CoV-2
- 46
- 47

48 Abstract

- 49 Genomic epidemiology offers important insight into the transmission and evolution of
- 50 respiratory viruses. We used metagenomic sequencing from negative SARS-CoV-2 antigen tests
- 51 to identify a wide range of respiratory viruses and generate full genome sequences, offering a
- 52 streamlined mechanism for broad respiratory virus genomic surveillance.

53 Introduction

54	The SARS-CoV-2 pandemic highlighted the importance of genomic epidemiology in
55	understanding virus transmission and evolution, informing essential countermeasures from non-
56	pharmaceutical interventions to vaccines. Massive global efforts in SARS-CoV-2 genomic
57	surveillance were made possible by widespread diagnostic testing and the growth of new
58	infrastructure and methods for sequencing and analysis (1). Most genomic surveillance pipelines
59	in the U.S. obtained residual SARS-CoV-2 positive samples from clinical, public health, and
60	commercial laboratories. This strategy was effective during the pandemic but difficult to
61	maintain with the rise of at-home rapid antigen tests (2, 3). As traditional sample sources
62	declined, our group and others demonstrated that residual samples from rapid antigen tests could
63	be used to generate and analyze full SARS-CoV-2 sequences for genomic surveillance (4-6).
64	Here, we build upon this work by identifying, sequencing, and analyzing other
65	respiratory viruses using residual swab samples from negative BinaxNow [™] COVID-19 antigen
66	tests. This multi-virus approach is important as SARS-CoV-2 has transitioned to an endemic
67	virus whose symptoms resemble those of other respiratory viruses (7). Thus, there is both a need
68	for broad testing and an opportunity to expand genomic surveillance for respiratory viruses using
69	self-collected samples.
70	

71 Methods

Detailed laboratory and analysis methods are provided in the Appendix. Briefly,
participants were enrolled in a parent study evaluating novel viral diagnostic tests through the
RADx program at the Atlanta Center for Microsystems Engineered Point-of-Care Technologies.
The study protocol was approved by the Emory Institutional Review Board and the Grady

76	Research Oversight Committee. We performed RNA metagenomic sequencing as described (8),
77	obtaining a median of 5.8 million reads per sample (Supplementary Data). We used a three-step
78	bioinformatic approach to detect viruses (Supplementary Figure 1) using KrakenUniq, blastn,
79	and reference mapping, with a final criterion requiring coverage of at least 3 distinct genome
80	regions, based on clinical diagnostic criteria for metagenomic sequencing (9).
81	
82	Results
83	We collected negative BinaxNOW TM test samples from 53 individuals between April-
84	August 2023 (Supplementary Table 1), a period during which 68% of the BinaxNOW TM tests in
85	the parent study were negative. All individuals were symptomatic at the time of testing (Table 1),
86	and the median interval between symptom onset and testing was 2 days (range 0-9). RT-PCR was
87	positive for influenza B in three samples and negative for influenza A and SARS-CoV-2 in all
88	samples (Supplementary Data).
89	Metagenomic sequencing identified a low level of SARS-CoV-2 in one sample and a
90	different pathogenic human respiratory virus in 17 of the other 52 samples (33%)
91	(Supplementary Data). The following viruses were detected: parainfluenza viruses (N=7),
92	rhinoviruses (N=5), influenza B (N=3), seasonal coronaviruses (N=2), and adenovirus (N=1)
93	(Figure 1). In one sample, both influenza B and parainfluenza 2 were detected. In another
94	sample positive for influenza B by RT-PCR, metagenomic sequencing did not identify influenza
95	but identified human mastadenovirus E. Thus, excluding SARS-CoV-2, a total of 18 viruses were
96	detected across 17 samples. There was no difference in the total number of reads obtained for
97	samples with and without viruses detected (Mann Whitney U test, p=0.29).

98	We observed potential differences in symptom frequencies between individuals with and
99	without viruses detected, but none were statistically significant (Table 1).

100 Of the 18 viruses detected, we generated full viral genome sequences from 11 (61%), 101 with >90% coverage and 71-24,000 fold depth (Supplementary Data). These included 102 parainfluenza 3 (4/4 samples), parainfluenza 2 (1/2), rhinovirus (5/5), and influenza B (1/3). 103 We performed phylogenetic analysis of parainfluenza 3 as a proof-of-concept for 104 genomic epidemiology studies and found substantial diversity. Using the lineage classification 105 system described in (10), two of our sequences clustered with Lineage A1 sequences from 2019-2023 (Figure 2A), another clustered with Lineage C sequences from Japan in 2023, and the 106 107 fourth with Lineage C sequences from the U.S. collected between 2015-2017 (Figure 2B), all 108 with high bootstrap support (Supplementary Figure 2). Of note, there are only about 450 109 complete parainfluenza 3 virus sequences available; the data from our small study represent 110 nearly 1% of this number, underscoring the opportunity to easily expand genomic surveillance 111 using this approach.

112 In addition to human pathogenic respiratory viruses, we detected over 100 viruses of no 113 clinical significance, including bacteriophages and plant viruses, many of which were also 114 detected in our negative controls (Figure 3). Similarly, mastadenovirus C was found in many 115 samples and negative controls. These are all consistent with environmental or reagent 116 contaminants. Herpesviruses were found in many samples by KrakenUniq and blastn, but were 117 not confirmed by mapping to a reference sequence with coverage of at least 3 regions. Overall, 118 1,367 viral taxa were identified by KrakenUniq, only 254 (18.6%) were confirmed by BLAST, 119 and only 137 (53.9% of these, 10% of total) met our criteria for detection, highlighting the 120 importance of confirmatory steps in metagenomic analysis.

121

122 Discussion

123	Our study demonstrates that RNA metagenomic sequencing of residual swab samples
124	from negative BinaxNOW TM tests can be used to detect a broad range of respiratory viruses,
125	including rhinoviruses, parainfluenza viruses, influenza B, seasonal coronaviruses, and
126	adenovirus. All of these have overlapping symptoms with one another and with SARS-CoV-2,
127	underscoring the need for multi-virus testing approaches. Although our study was not designed
128	for clinical diagnosis, metagenomic sequencing is increasingly used clinically, and our results
129	illustrate the need for rigorous analysis techniques and careful interpretation.
130	It is notable that only 33% of samples had a human pathogenic respiratory virus. This is
131	similar to our prior study detecting alternative respiratory viruses in only 40% of SARS-CoV-2
132	negative individuals using residual clinical samples early in the pandemic (8). Possible
133	explanations include individuals with a non-infectious syndrome, a bacterial or other non-viral
134	infection, or a virus present at a low level. It is also possible that some individuals were infected
135	with a DNA virus not optimally captured by RNA sequencing. However, we detected adenovirus,
136	the most prevalent respiratory DNA virus. Among common RNA viruses, we did not detect
137	influenza A or RSV, which we attribute to the winter-predominant seasonality of these viruses
138	compared to our sample collection in spring and summer.
139	Importantly, of the 18 viruses detected, we were able to generate full viral genome
140	sequences from 11 (61%) using moderate sequencing depths. Thus, the single laboratory
141	technique of metagenomic sequencing can not only identify diverse respiratory viruses but also
142	contribute to their genomic surveillance. The surprisingly high depth of genome coverage

achieved for many sequences indicates that throughput and cost can be improved by reducingtotal sequencing reads from each sample in future studies.

145	By combining metagenomic sequencing with the use of residual antigen test samples, we
146	demonstrate a mechanism for convenient and broad respiratory virus surveillance. Our study
147	used BinaxNOW TM tests, which conveniently preserve the used swab within the kit cassette;
148	future work is needed to evaluate this approach using rapid antigen test strips themselves, as
149	previously demonstrated for SARS-CoV-2 sequencing (5). Additionally, future studies would
150	benefit from a regulatory framework in which results can be returned to study participants, who
151	are likely curious about the presence of other respiratory viruses when rapid antigen testing is
152	negative.
153	In conclusion, our study illustrates that residual samples from self-collected antigen tests
154	can be a powerful sample source for investigating the genomic epidemiology of a broad range of
155	respiratory viruses, building upon the strong foundations for viral surveillance established during
156	the SARS-CoV-2 pandemic.
157	
158	Data Availability:
159	All raw sequencing data (cleaned of human reads) is available in NCBI SRA under BioProject
160	PRJNA1144955, and assembled virus genome sequences are available in NCBI GenBank with
161	accession numbers listed in the Supplementary Data file.
162	
163	Acknowledgements:
164	We would like to thank the study participants.
165	

166 Disclosures:

167 All authors report no conflicts of interest to disclose.

168

169 Funding:

- 170 This work was supported by NIH U54 EB027690 02S1, U54 EB027690 03S1, U54EB027690
- 171 03S2 UL1 TR002378 and the Centers for Disease Control and Prevention-funded Georgia
- 172 Pathogen Genomics Center of Excellence contract 40500-050-23234506. This study was
- supported in part by the Emory Integrated Genomics Core (EIGC) (RRID:SCR_023529), which
- is subsidized by the Emory University School of Medicine and is one of the Emory Integrated
- 175 Core Facilities. Additional support was provided by the Georgia Clinical & Translational Science
- 176 Alliance of the National Institutes of Health under Award Number UL1TR002378. The content is
- solely the responsibility of the authors and does not necessarily reflect the official views of the
- 178 National Institutes of Health.

179

180 Author bio:

181 Ms. Jules received a Bachelor of Science in Anthropology and Human Biology from Emory

182 University and is currently a research specialist in the Department of Pathology and Laboratory

183 Medicine in the Emory University School of Medicine. She will be applying to medical school

184 with the aspiration of becoming a family doctor and expanding healthcare to underserved

185 communities.

186 **References**

- 187 1. Oude Munnink BB, Worp N, Nieuwenhuijse DF, Sikkema RS, Haagmans B, Fouchier
- 188 RAM, et al. The next phase of SARS-CoV-2 surveillance: real-time molecular epidemiology. Nat
- 189 Med. 2021;27(9):1518-24. Epub 20210909. doi: 10.1038/s41591-021-01472-w. PubMed PMID:
- 190 34504335.
- 191 2. Ritchey MD, Rosenblum HG, Del Guercio K, Humbard M, Santos S, Hall J, et al.
- 192 COVID-19 Self-Test Data: Challenges and Opportunities United States, October 31, 2021-June
- 193 11, 2022. MMWR Morb Mortal Wkly Rep. 2022;71(32):1005-10. Epub 20220812. doi:
- 194 10.15585/mmwr.mm7132a1. PubMed PMID: 35951486; PubMed Central PMCID:
- 195 PMC9400539.
- 196 3. Rader B, Gertz A, Iuliano AD, Gilmer M, Wronski L, Astley CM, et al. Use of At-Home
- 197 COVID-19 Tests United States, August 23, 2021-March 12, 2022. MMWR Morb Mortal Wkly
- 198 Rep. 2022;71(13):489-94. Epub 20220401. doi: 10.15585/mmwr.mm7113e1. PubMed PMID:
- 199 35358168; PubMed Central PMCID: PMC8979595.
- 200 4. Nguyen PV, Carmola LR, Wang E, Bassit L, Rao A, Greenleaf M, et al. SARS-CoV-2
- 201 molecular testing and whole genome sequencing following RNA recovery from used BinaxNOW
- 202 COVID-19 antigen self tests. J Clin Virol. 2023;162:105426. Epub 20230324. doi:
- 203 10.1016/j.jcv.2023.105426. PubMed PMID: 37028004; PubMed Central PMCID:
- 204 PMC10036152.
- 205 5. Martin GE, Taiaroa G, Taouk ML, Savic I, O'Keefe J, Quach R, et al. Maintaining
- 206 genomic surveillance using whole-genome sequencing of SARS-CoV-2 from rapid antigen test
- 207 devices. Lancet Infect Dis. 2022;22(10):1417-8. Epub 20220804. doi: 10.1016/S1473-
- 208 3099(22)00512-6. PubMed PMID: 35934015; PubMed Central PMCID: PMC9352270.

209	6.	Hassouneh SA,	Trujillo A, Ali S,	Cella E, Johnston	C, DeRuff KC,	, et al. Antigen test
-----	----	---------------	--------------------	-------------------	---------------	-----------------------

- swabs are comparable to nasopharyngeal swabs for sequencing of SARS-CoV-2. Sci Rep.
- 211 2023;13(1):11255. Epub 20230712. doi: 10.1038/s41598-023-37893-5. PubMed PMID:
- 212 37438412; PubMed Central PMCID: PMC10338537.
- 213 7. Geismar C, Nguyen V, Fragaszy E, Shrotri M, Navaratnam AMD, Beale S, et al.
- 214 Symptom profiles of community cases infected by influenza, RSV, rhinovirus, seasonal
- coronavirus, and SARS-CoV-2 variants of concern. Sci Rep. 2023;13(1):12511. Epub 20230802.
- 216 doi: 10.1038/s41598-023-38869-1. PubMed PMID: 37532756; PubMed Central PMCID:
- 217 PMC10397315.
- 218 8. Babiker A, Bradley HL, Stittleburg VD, Ingersoll JM, Key A, Kraft CS, et al.
- 219 Metagenomic Sequencing To Detect Respiratory Viruses in Persons under Investigation for
- 220 COVID-19. J Clin Microbiol. 2020;59(1). Epub 20201217. doi: 10.1128/JCM.02142-20.
- 221 PubMed PMID: 33067271; PubMed Central PMCID: PMC7771462.
- 9. Miller S, Naccache SN, Samayoa E, Messacar K, Arevalo S, Federman S, et al.
- 223 Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in
- 224 cerebrospinal fluid. Genome Res. 2019;29(5):831-42. Epub 20190416. doi:
- 225 10.1101/gr.238170.118. PubMed PMID: 30992304; PubMed Central PMCID: PMC6499319.
- 226 10. Lee K, Park K, Sung H, Kim MN. Phylogenetic lineage dynamics of global parainfluenza
- virus type 3 post-COVID-19 pandemic. mSphere. 2024;9(4):e0062423. Epub 20240319. doi:
- 228 10.1128/msphere.00624-23. PubMed PMID: 38501829; PubMed Central PMCID:
- 229 PMC11036794.

230

231 Tables:

232

- 233 Table 1: Participant symptoms. Table indicates the number and percent of participants
- reporting each symptom at the time of testing. For symptom categories (gray rows), the number
- of participants with at least one symptom in that category is reported. *This includes one
- individual with SARS-CoV-2 detected at a low level and 17 individuals with an alternative
- 237 human pathogenic respiratory virus detected.

Symptom	Total participants (N = 53)	Participants with a virus detected (N = 18*)	Participants with no virus detected (N = 35)
Upper Respiratory	47 (88.7%)	17 (94.4%)	30 (85.7%)
Congestion/runny nose	37 (69.8%)	15 (83.3%)	22 (62.9%)
Sore throat Loss of sense of taste or	33 (62.3%)	14 (77.8%)	19 (54.3%)
smell	7 (13.2%)	2 (11.1%)	5 (14.3%)
Lower Respiratory	43 (81.1%)	15 (83.3%)	28 (80.0%)
Cough	39 (73.6%)	15 (83.3%)	24 (68.6%)
Shortness of breath	23 (43.4%)	6 (33.3%)	17 (48.6%)
Gastrointestinal	15 (28.3%)	6 (33.3%)	9 (25.7%)
Vomiting	4 (7.6%)	3 (16.7%)	1 (2.9%)
Nausea	11 (20.8%)	2 (11.1%)	9 (25.7%)
Diarrhea	2 (3.8%)	1 (5.6%)	1 (2.9%)
Abdominal pain	7 (13.2%)	2 (11.1%)	5 (14.3%)
Systemic	35 (66.0%)	13 (72.2%)	22 (62.9%)
Fever (>100.4)	18 (34.0%)	7 (38.9%)	11 (31.4%)
Chills	24 (45.3%)	11 (61.1%)	13 (37.1%)
Fatigue	27 (50.9%)	11 (61.1%)	16 (45.71%)
Other	41 (77.4%)	16 (88.9%)	25 (71.4%)
Headache	30 (56.6%)	13 (72.2%)	17 (48.6%)
Joint pain	14 (26.4%)	3 (16.7%)	11 (31.4%)
Muscle pain	31 (58.5%)	10 (55.6%)	21 (60.0%)

238

239 **Figures:**

210

240 241	Figure 1: Frequency of human pathogenic respiratory viruses found in 53 residual samples
242	from BinaxNOW TM tests that were negative for SARS-CoV-2. Pie charts indicate the number
243	of samples positive for each virus among all samples (left panel) and among the 18 positive
244	samples (right panel). Numbers indicate the number of samples with a virus identified, followed
245	in parentheses by the number of samples with a $>90\%$ complete genome sequence assembled.
246	
247	Figure 2: Phylogenetic analysis of parainfluenza 3 virus sequences. The names of sequences
248	obtained in this study are bold and in red, and reference sequences are in black. The outer ring
249	indicates virus lineage: A) contains representative lineage A1 sequences, and B) contains
250	representative sequences from lineages C, E, F, and G. Each tree is a maximized parsimony
251	subtree using down sampled data from the full analysis in Supplementary Figure 2, for ease of
252	visualization.
253	
254	Figure 3: Plot of the viral taxa (rows) that were detected in each sample (columns). Blue
255	boxes indicate viruses that were detected by both KrakenUniq and blastn but not confirmed by
256	reference mapping, while red boxes indicate viruses that were detected by both KrakenUniq and
257	blastn and were confirmed by reference mapping (covering at least 3 distinct regions of the
258	reference virus genome).
259	
260	Supplementary Figure 1: Analysis pipeline for metagenomic classification and
261	confirmation. Red boxes describe processing and quality control steps, yellow boxes describe

262 initial metagenomic classification using KrakenUniq, green boxes describe blastn confirmation

- 263 of reads classified as viral by KrakenUniq, and blue boxes describe reference-based mapping for
- final confirmation. LCA = lowest common ancestor.

265

266 Supplementary Figure 2: Maximum likelihood phylogenetic analysis of parainfluenza 3

- 267 virus sequences. The names of sequences obtained in this study are bold and in red, and
- 268 reference sequences in black represent all unique full-length genome sequences of parainfluenza
- 269 3 available in GenBank (7/30/24). Circles indicate nodes with >95% ultrafast bootstrap support.
- 270 The outer ring indicates virus lineage.



Virus-positive samples













Legend Confirmed

Not confirmed





Bacteriophages

