

1 **Detection of Mutations Associated with Variants of Concern Via High Throughput**  
2 **Sequencing of SARS-CoV-2 Isolated from NYC Wastewater**

3

4 Davida S. Smyth<sup>a,\*</sup>, Monica Trujillo<sup>b,\*</sup>, Kristen Cheung<sup>c,¶</sup>, Anna Gao<sup>c,¶</sup>, Irene Hoxie<sup>c,d,¶</sup>,  
5 Sherin Kannoly<sup>c,¶</sup>, Nanami Kubota<sup>c,¶</sup>, Michelle Markman<sup>c,¶</sup>, Kaung Myat San<sup>c,¶</sup>, Geena  
6 Sompanya<sup>a,¶</sup>, and John J. Dennehy<sup>c,d\*</sup>

7

8 <sup>a</sup>Department of Natural Sciences and Mathematics, Eugene Lang College of Liberal  
9 Arts at The New School, New York

10 <sup>b</sup>Department of Biology, Queensborough Community College, The City University of  
11 New York, New York

12 <sup>c</sup>Biology Department, Queens College, The City University of New York, New York

13 <sup>d</sup>Biology Doctoral Program, The Graduate Center, The City University of New York, New  
14 York

15

16 Running title: Variants of Concern Detected in NYC Wastewater (47 characters)

17 \*Joint first authors

18 ¶Authors in alphabetical order

19 #Address correspondence to: [john.dennehy@qc.cuny.edu](mailto:john.dennehy@qc.cuny.edu)

20

21 **ABSTRACT (186 words)**

22 Monitoring SARS-CoV-2 genetic diversity is strongly indicated because diversifying  
23 selection may lead to the emergence of novel variants resistant to naturally acquired or  
24 vaccine-induced immunity. To date, most data on SARS-CoV-2 genetic diversity has  
25 come from the sequencing of clinical samples, but such studies may suffer limitations  
26 due to costs and throughput. Wastewater-based epidemiology may provide an  
27 alternative and complementary approach for monitoring communities for novel variants.  
28 Given that SARS-CoV-2 can infect the cells of the human gut and is found in high  
29 concentrations in feces, wastewater may be a valuable source of SARS-CoV-2 RNA,  
30 which can be deep sequenced to provide information on the circulating variants in a  
31 community. Here we describe a safe, affordable protocol for the sequencing of SARS-  
32 CoV-2 RNA using high-throughput Illumina sequencing technology. Our targeted  
33 sequencing approach revealed the presence of mutations associated with several  
34 Variants of Concern at appreciable frequencies. Our work demonstrates that  
35 wastewater-based SARS-CoV-2 sequencing can inform surveillance efforts monitoring  
36 the community spread of SARS-CoV-2 Variants of Concern and detect the appearance  
37 of novel emerging variants more cheaply, safely, and efficiently than the sequencing of  
38 individual clinical samples.

39 **IMPORTANCE (140 words)**

40 The SARS-CoV-2 pandemic has caused millions of deaths around the world as  
41 countries struggle to contain infections. The pandemic will not end until herd immunity is  
42 reached, that is, when most of the population has either recovered from SARS-CoV-2  
43 infection or is vaccinated against SARS-CoV-2. However, the emergence of new SARS-  
44 CoV-2 variants of concern threatens to erase gains. Emerging new variants may re-  
45 infect persons who have recovered from COVID-19 or may evade vaccine-induced  
46 immunity. However, scaling up SARS-CoV-2 genetic sequencing to monitor Variants of  
47 Concern in communities around the world is challenging. Wastewater-based  
48 sequencing of SARS-CoV-2 RNA can be used to monitor the presence of emerging  
49 variants in large communities to enact control measures to minimize the spread of these  
50 variants. We describe here the identification of alleles associated with several variants  
51 of concern in wastewater obtained from NYC watersheds.

52 **KEYWORDS:** coronavirus, environmental microbiology, Illumina sequencing,  
53 metagenomics, NGS, sewage, virus surveillance, Variants of Concern, wastewater-  
54 based epidemiology

55 **INTRODUCTION**

56 The emergence of novel SARS-CoV-2 Variants of Concern, including B.1.1.7 from the  
57 United Kingdom and B.1.351 from South Africa, has provoked intense speculation about  
58 the future of the pandemic (1-3). Early studies suggest that these new variants may be  
59 more transmissible (4-6). Even more concerning are reports of decreased antibody-  
60 mediated neutralization of these variants (7-9). Regardless of the biological attributes of

61 these novel variants, it is clear that behavioral interventions, public health measures,  
62 vaccinations, and reduced numbers of susceptible individuals will impose strong  
63 diversifying selection on SARS-CoV-2 to enhance transmission and/or evade host  
64 immunity (10). We should anticipate that the continued evolution of SARS-CoV-2 may  
65 result in variants that evade natural or vaccine-mediated immunity. As such, intensive  
66 monitoring of SARS-CoV-2 genetic diversity and evolution is vital to rapidly identify  
67 Variants of Concern as they emerge.

68 Currently, most SARS-CoV-2 genetic surveillance is conducted via the genome  
69 sequencing of viral RNA obtained from clinical specimens. While occurring at a much  
70 greater rate and volume than previous epidemics, the sequencing of clinical specimens  
71 is limited by cost, coverage, quality, and throughput concerns. In developed countries,  
72 these issues are not readily apparent, but sequencing efforts in underdeveloped  
73 countries has been more restricted (11). Another disadvantage of focusing on clinical  
74 strains stems from the large number of asymptomatic or mildly symptomatic infections  
75 (12). SARS-CoV-2 sequencing efforts will suffer biases if genomic information is more  
76 frequently obtained from seriously ill patients, rather than from asymptomatic patients,  
77 and those with mild symptoms who choose to follow the CDC's advice and convalesce  
78 at home. Wastewater-based epidemiology may provide an alternative and  
79 complementary approach to provide more representative SARS-CoV-2 genetic data at  
80 lower costs and higher throughput.

81 Given that SARS-CoV-2 has been detected in fecal samples (13, 14), and subsequently  
82 in wastewater, wastewater is being monitored in communities around the world to  
83 determine SARS-CoV-2 prevalence in communities (15-17). Furthermore, isolation of  
84 SARS-CoV-2 RNA from wastewater coupled with high-throughput deep sequencing  
85 provides an almost unlimited source of unbiased viral sequences, which can be used to  
86 monitor frequencies of Variants of Concern in populations (18-20). We have focused on  
87 the use of targeted sequencing of the spike genomic region known to encode Variants  
88 of Concern. Our approach, while limited to a specific region of the genome, is  
89 affordable, rapid and generates sufficient coverage to quantify known variants and to  
90 identify possible emerging ones.

91 Our team, in conjunction with the New York City Department of Environmental  
92 Protection, has been monitoring the genetic signal of SARS-CoV-2 in the wastewater of  
93 all 14 wastewater treatment plants in NYC, an area that encompasses a population of  
94 8,419,000 persons, since June 2020. We developed and optimized a protocol for safe,  
95 cost-effective, and repeatable quantitation of SARS-CoV-2 copy number by RT-qPCR  
96 (21). Our protocol performed strongly in a large-scale, nationwide comparative study of  
97 the reproducibility and sensitivity of 36 methods of quantifying SARS-CoV-2 in  
98 wastewater (22). Our protocol is identified as 4S.1(H) in Table 3. We further extended  
99 the utility of our protocol by deep sequencing SARS-CoV-2 RNA isolated from  
100 wastewater samples. Here we report presence of alleles associated with different  
101 Variants of Concern at appreciable frequencies. Our findings provide support for recent

102 observations of increasing frequencies of New York Variant of Interest B.1.526 in  
103 clinical samples (23, 24), as well as the presence of Variants of Concern from the  
104 United Kingdom, California, South Africa and Brazil (25). Furthermore, our results  
105 demonstrate the utility of wastewater-based epidemiology for the timely identification of  
106 novel variants of concern arising in communities.

## 107 **RESULTS AND DISCUSSION**

### 108 **Targeted sequencing is a viable approach for identifying SARS-CoV-2 mutations.**

109 We generated cDNA from NYC wastewater samples that exhibited RT-qPCR Cts values  
110 ranging from 28 to 24 Cts corresponding to 26,443 and 1,423,339 N1 copies/L,  
111 respectively. Using this cDNA as a template, we PCR amplified a region of the receptor  
112 binding domain (RBD) of the SARS-CoV-2 Spike gene, spanning amino acid residues  
113 P410 to L513, which encompasses mutations that are found in several known Variants  
114 of Concern. A total of 420 single nucleotide variants were identified in the 45 samples  
115 sequenced (Supplementary Table 1). Coverage ranged from 1,037x – 118,737x with a  
116 mean of 23,586x (Supplementary Table 1). Across all samples, we identified 75 unique  
117 mutations resulting in amino acid substitutions, 20 unique synonymous mutations, and  
118 18 deletions resulting in a frameshift, in the 332 bp region targeted (Supplementary  
119 Table 1).

### 120 **Mutations associated with Variants of Concern are present in NYC wastewater.**

121 The five mutations found at highest frequencies, both in terms of frequency of reads  
122 within samples and found in the most samples, were L452R, E484K, N501Y, S494P,  
123 and S477N. All five mutations are associated with known Variants of Concern (Fig. 1;  
124 Supplementary Table 2). On Jan 31<sup>st</sup>, we sequenced samples from two wastewater  
125 treatment plants in NYC and identified reads containing mutations L452R, S477N,  
126 E484K, S494P and N501Y in both. On February 28<sup>th</sup> and March 14<sup>th</sup> samples from all  
127 14 wastewater treatment plants in NYC were sequenced, revealing the presence of a  
128 high proportion of reads containing mutations L452R, S477N, E484K, S494P and  
129 N501Y (Fig. 1). Mutation L452R is unique to Pango lineage Variants of Concern  
130 B.1.427 and B.1.429, which were first observed in California (25, 26). Mutation S477N is  
131 only found in New York Variant of Interest B.1.526 (23-25, 27). Mutation E484K has  
132 been reported in Variants of Concern B.1.1.7 from the United Kingdom, P.1 and P.2  
133 from Brazil, and B.1.351 from South Africa, and B.1.525 and B.1.526 from New York  
134 (25). Mutation S494P is only found in Variant of Concern B.1.1.7 from the United  
135 Kingdom (25). Mutation N501Y is found in Variants of Concern B.1.1.7 from the United  
136 Kingdom, P.1 from Brazil, and B.1.351 from South Africa (25).

137 The finding that unique mutations associated with different Variants of Concern in our  
138 pooled sequencing assay suggests the circulation of these variants in NYC. A caveat  
139 with our approach, however, is that we cannot conclusively identify the presence of a  
140 Variant of Concern since our sequencing assay targets only a region of the receptor  
141 binding domain, and some significant mutations are outside the sequenced region.  
142 Furthermore, additional mutations occurring in the primer binding region may allow

143 some mutations to go undetected because their DNA could not be amplified. We are  
144 expanding our targeted sequencing approach to include additional regions of interest to  
145 minimize the chance of missing important variants. Additionally, we intend to generate  
146 cDNA with random hexamers, and to incorporate a level of degeneracy in the  
147 sequencing primers to increase the breadth of our targeted sequencing.

148 Our most recent data from March 14<sup>th</sup> suggests a slight decrease in the prevalence of  
149 the E484K variant, but we cannot draw firm conclusions due to the nature of our  
150 sequencing assay, which relies on the collective sequencing of a large pool of  
151 individuals. Nevertheless, our frequency data agrees with that recently observed in  
152 human clinical samples from NYC (23, 24, 27). We intend to supplement our targeted  
153 sequencing approach with whole genome amplicon sequencing in the future.

154 We believe that our approach offers a viable alternative to whole genome sequencing  
155 for the detection of known variants and can be rapidly deployed to detect additional  
156 emerging variants of concern. Importantly as a cost saving measure, labs can generate  
157 the libraries themselves and outsource the sequencing component to companies/core  
158 facilities if they lack access to a sequencer, generally with a short turnaround time.

## 159 **MATERIALS AND METHODS**

160 **Wastewater Sample Processing and RNA Extraction.** Wastewater was collected  
161 from 14 NYC wastewater treatment plants and RNA isolated according to our previously  
162 published protocol ([dx.doi.org/10.17504/protocols.io.brr6m59e](https://doi.org/10.17504/protocols.io.brr6m59e)) (21). Control SARS-  
163 CoV-2 synthetic RNA was purchased from Twist Bioscience (#102019).

164 Briefly, 250 mL from 24-hr composite raw sewage samples were obtained from NYC  
165 wastewater treatment plants (WWTPs) and centrifuged at 5,000 x g for 10 min at 4°C to  
166 pellet solids. 40 mL of supernatant was passed through a 0.22 µM filter. Filtrate was  
167 stored at 4°C for 24 hrs after adding 0.9 g sodium chloride and 4.0 g PEG 8000 (Fisher)  
168 then centrifuged at 12,000 x g for 120 minutes at 4 °C to pellet precipitate. The pellet  
169 was resuspended in 1.5 mL TRIzol (Fisher), and RNA was purified according to the  
170 manufacturer's instructions.

171 **Targeted PCR.** Our target for sequencing was a 332 bp region of the Receptor Binding  
172 Domain (RBD) of the spike protein spanning amino acid residues P420 to L513.

173 Mutations in this region are of critical importance as they might help the variants evade  
174 current antibody treatments and vaccines. RNA isolated from wastewater was used to  
175 generate cDNA using ProtoScript® II Reverse Transcriptase (New England Biolabs).

176 The RNA was incubated with an RBD specific primer (ccagatgattttacaggctgcg) and  
177 dNTPs (0.5 mM final concentration) at 65°C for 5 minutes and placed on ice. The RT  
178 buffer, DTT (0.01 M final concentration), and the RT were added to the same tube and  
179 incubated at 42°C for 2 hours followed by 20 minutes at 65°C to inactivate the enzyme.  
180 The RBD region was amplified using Q5® High-Fidelity DNA Polymerase using the  
181 forward primer 5' -

182 TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGccagatgattttacaggctgcg-3' and

183 reverse primer 5'-  
184 GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGgaaagtactactactctgtatggtgg-3',  
185 which incorporate Illumina adaptors. PCR performed as follows: 98°C for 30 seconds,  
186 followed by 40 cycles of 98°C 5 seconds, 53°C for 15 seconds and 65°C for 1 minute and  
187 a final extension at 65°C for 1 minute.

188 **Targeted Sequencing.** The RBD amplicons were purified using AMPure XP beads  
189 (Beckman Coulter). Index PCR was performed using the Nextera DNA CD Indexes kit  
190 (Illumina) with 2X KAPA HiFi HotStart ReadyMix (Roche), and indexed PCR products  
191 purified using AMPure beads. The indexed libraries were quantified using the Qubit 3.0  
192 and Qubit dsDNA HS Assay Kit and diluted in 10 mM Tris-HCl to a final concentration of  
193 approximately 0.3 ng/μL (1 nM). The libraries were pooled together and diluted to a final  
194 concentration of 50 pM. Before sequencing on an Illumina iSeq100, a 10% spike-in of  
195 50 pM PhiX control v3 (Illumina) was added to the pooled library.

196 **Bioinformatics.** Sequencing data was uploaded to the BaseSpace Sequence Hub, and  
197 the reads demultiplexed using a FASTQ generation script. Reads were processed using  
198 the published Geneious workflows for preprocessing of NGS reads and assembly of  
199 SARS-CoV-2 amplicons ([https://help.geneious.com/hc/en-us/articles/360045070991-  
200 Assembly-of-SARS-CoV-2-genomes-from-tiled-amplicon-Illumina-sequencing-using-  
201 Geneious-Prime](https://help.geneious.com/hc/en-us/articles/360045070991-Assembly-of-SARS-CoV-2-genomes-from-tiled-amplicon-Illumina-sequencing-using-Geneious-Prime) and [https://help.geneious.com/hc/en-us/articles/360044626852-Best-  
202 practice-for-preprocessing-NGS-reads-in-Geneious-Prime](https://help.geneious.com/hc/en-us/articles/360044626852-Best-practice-for-preprocessing-NGS-reads-in-Geneious-Prime)). Paired reads were trimmed,  
203 and the adapter sequences removed with the BBDuk plugin. Trimmed reads were  
204 merged and aligned to the SARS-CoV-2 reference genome MN908947. Variants were  
205 called using the Annotate and Predict Find Variations/SNPs in Geneious and verified by  
206 using the V-PIPE SARS-CoV-2 application ([https://cbg-ethz.github.io/V-pipe/sars-cov-  
207 2/](https://cbg-ethz.github.io/V-pipe/sars-cov-2/))(28).

## 208 **Data Availability**

209 Raw sequencing reads are available in NCBI's Sequence Read Archive (SRA) under  
210 accession # PRJNA715712.

## 211 **ACKNOWLEDGEMENTS**

212 The research described herein would not be possible if not for the assistance and  
213 support of a wide-range of organizations and individuals that came together to address  
214 the shared calamity that is the COVID-19 pandemic. We thank Gina Behnke, Esmeraldo  
215 Castro, Francoise Chauvin, Alexander Clare, Pilar Domingo-Calap, Pam Elardo, Raul  
216 Gonzalez, Crystal Hepp, Catherine Hoar, Dimitrios Katehis, William Kelly, Samantha  
217 McBride, Hope McGibbon, Hilary Millar, Samantha Patinella, Krish Ramalingam, Andrea  
218 Silverman, Jasmin Torres, Arvind Varsani, Peter Williamsen for support, advice,  
219 discussions, and feedback. This work was funded in part by the New York City  
220 Department of Environmental Protection. The Water Research Foundation, the NSF  
221 Research Coordination Network for Wastewater Surveillance for SARS-CoV-2 and

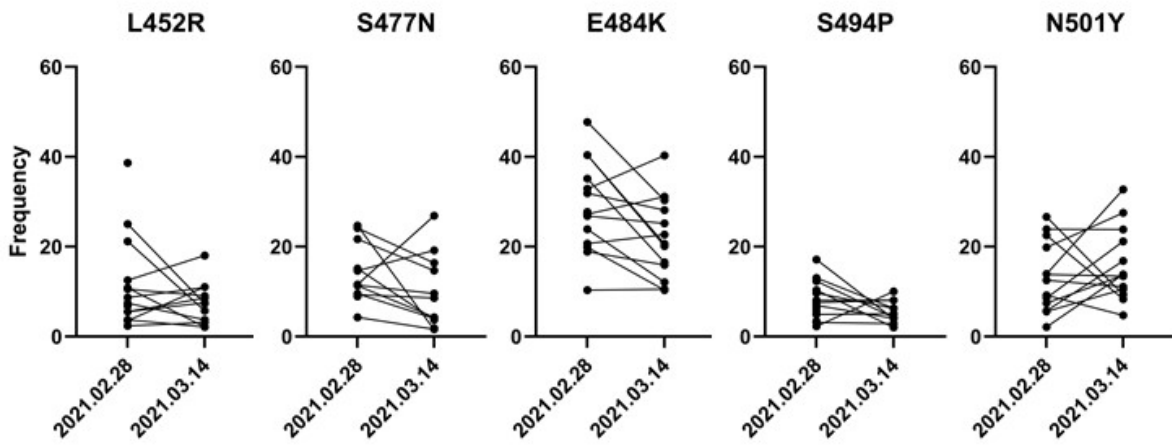
222 Qiagen Inc. provided resources, materials and supplies, technical support, and  
223 community support.

224 **FIGURE LEGEND**

225 **Figure 1.** Frequencies of reads associated with five selected mutations associated with  
226 SARS-CoV-2 Variants of Concern from wastewater obtained from 14 NYC wastewater  
227 treatment plants on two separate dates.

228

229 **Figure 1**



230

231

232



233 **REFERENCES**

234 1. Alpert T, Lasek-Nesselquist E, Brito AF, Valesano AL, Rothman J, MacKay MJ, Petrone ME,  
235 Breban MI, Watkins AE, Vogels CBF, Russell A, Kelly JP, Shudt M, Plitnick J, Schneider E,  
236 Fitzsimmons WJ, Khullar G, Metti J, Dudley JT, Nash M, Wang J, Liu C, Hui P, Muyombwe A,  
237 Downing R, Razeq J, Bart SM, Murphy S, Neal C, Laszlo E, Landry ML, Cook PW, Fauver JR, Mason  
238 CE, Lauring AS, St George K, MacCannell DR, Grubaugh ND. 2021. Early introductions and  
239 community transmission of SARS-CoV-2 variant B.1.1.7 in the United States. medRxiv : the  
240 preprint server for health sciences doi:10.1101/2021.02.10.21251540:2021.02.10.21251540.

241 2. Washington NL, Gangavarapu K, Zeller M, Bolze A, Cirulli ET, Schiabor Barrett KM, Larsen BB,  
242 Anderson C, White S, Cassens T, Jacobs S, Levan G, Nguyen J, Ramirez JM, Rivera-Garcia C,  
243 Sandoval E, Wang X, Wong D, Spencer E, Robles-Sikisaka R, Kurzban E, Hughes LD, Deng X, Wang  
244 C, Servellita V, Valentine H, De Hoff P, Seaver P, Sathe S, Gietzen K, Sickler B, Antico J, Hoon K,  
245 Liu J, Harding A, Bakhtar O, Basler T, Austin B, Isaksson M, Febbo P, Becker D, Laurent M,  
246 McDonald E, Yeo GW, Knight R, Laurent LC, de Feo E, Worobey M, Chiu C, Suchard MA, et al.  
247 2021. Genomic epidemiology identifies emergence and rapid transmission of SARS-CoV-2 B.1.1.7  
248 in the United States. medRxiv : the preprint server for health sciences  
249 doi:10.1101/2021.02.06.21251159:2021.02.06.21251159.

250 3. Leung K, Shum MH, Leung GM, Lam TT, Wu JT. 2021. Early transmissibility assessment of the  
251 N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. Euro  
252 surveillance : bulletin Europeen sur les maladies transmissibles = European communicable  
253 disease bulletin 26:2002106.

254 4. Zhao S, Lou J, Cao L, Zheng H, Chong MKC, Chen Z, Chan RWY, Zee BCY, Chan PKS, Wang MH.  
255 2021. Quantifying the transmission advantage associated with N501Y substitution of SARS-CoV-  
256 2 in the UK: an early data-driven analysis. J Travel Med 28.

257 5. Hunter PR, Brainard J, Grant A. 2021. The Impact of the November 2020 English National  
258 Lockdown on COVID-19 case counts. medRxiv  
259 doi:10.1101/2021.01.03.21249169:2021.01.03.21249169.

260 6. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, Hinsley WR, Laydon DJ, Dabrera  
261 G, O'Toole Á, Amato R, Ragonnet-Cronin M, Harrison I, Jackson B, Ariani CV, Boyd O, Loman NJ,  
262 McCrone JT, Gonçalves S, Jorgensen D, Myers R, Hill V, Jackson DK, Gaythorpe K, Groves N,  
263 Sillitoe J, Kwiatkowski DP, Flaxman S, Ratmann O, Bhatt S, Hopkins S, Gandy A, Rambaut A,  
264 Ferguson NM. 2021. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from  
265 linking epidemiological and genetic data. medRxiv  
266 doi:10.1101/2020.12.30.20249034:2020.12.30.20249034.

267 7. Collier DA, De Marco A, Ferreira I, Meng B, Datir R, Walls AC, Kemp SS, Bassi J, Pinto D, Fregni CS,  
268 Bianchi S, Tortorici MA, Bowen J, Culap K, Jaconi S, Cameroni E, Snell G, Pizzuto MS, Pellanda AF,  
269 Garzoni C, Riva A, Elmer A, Kingston N, Graves B, McCoy LE, Smith KG, Bradley JR, Temperton N,  
270 Ceron-Gutierrez LL, Barcenas-Morales G, Harvey W, Virgin HW, Lanzavecchia A, Piccoli L,  
271 Doffinger R, Wills M, Veesler D, Corti D, Gupta RK. 2021. SARS-CoV-2 B.1.1.7 sensitivity to mRNA  
272 vaccine-elicited, convalescent and monoclonal antibodies. medRxiv  
273 doi:10.1101/2021.01.19.21249840.

274 8. Tada T, Dcosta BM, Samanovic-Golden M, Herati RS, Cornelius A, Mulligan MJ, Landau NR. 2021.  
275 Neutralization of viruses with European, South African, and United States SARS-CoV-2 variant  
276 spike proteins by convalescent sera and BNT162b2 mRNA vaccine-elicited antibodies. bioRxiv  
277 doi:10.1101/2021.02.05.430003.

278 9. Graham C, Seow J, Huettner I, Khan H, Kouphou N, Acors S, Winstone H, Pickering S, Pedro  
279 Galao R, Jose Lista M, Jimenez-Guardeno JM, Laing AG, Wu Y, Joseph M, Muir L, Ng WM,

280 Duyvesteyn HME, Zhao Y, Bowden TA, Shankar-Hari M, Rosa A, Cherepanov P, McCoy LE,  
281 Hayday AC, Neil SJD, Malim MH, Doores KJ. 2021. Impact of the B.1.1.7 variant on neutralizing  
282 monoclonal antibodies recognizing diverse epitopes on SARS-CoV-2 Spike. bioRxiv  
283 doi:10.1101/2021.02.03.429355.

284 10. Dennehy JJ. 2017. Evolutionary ecology of virus emergence. *Annals of the New York Academy of*  
285 *Sciences* 1389:124-146.

286 11. Furuse Y. 2021. Genomic sequencing effort for SARS-CoV-2 by country during the pandemic. *Int J*  
287 *Infect Dis* 103:305-307.

288 12. Johansson MA, Quandelacy TM, Kada S, Prasad PV, Steele M, Brooks JT, Slayton RB, Biggerstaff  
289 M, Butler JC. 2021. SARS-CoV-2 Transmission From People Without COVID-19 Symptoms. *JAMA*  
290 *Network Open* 4:e2035057-e2035057.

291 13. Chen Y, Chen L, Deng Q, Zhang G, Wu K, Ni L, Yang Y, Liu B, Wang W, Wei C, Yang J, Ye G, Cheng  
292 Z. 2020. The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. *J Med Virol* 92:833-  
293 840.

294 14. Walsh KA, Jordan K, Clyne B, Rohde D, Drummond L, Byrne P, Ahern S, Carty PG, O'Brien KK,  
295 O'Murchu E, O'Neill M, Smith SM, Ryan M, Harrington P. 2020. SARS-CoV-2 detection, viral load  
296 and infectivity over the course of an infection. *J Infect* 81:357-371.

297 15. Larsen DA, Wigginton KR. 2020. Tracking COVID-19 with wastewater. *Nature Biotechnology*  
298 38:1151-1153.

299 16. Medema G, Been F, Heijnen L, Petterson S. 2020. Implementation of environmental surveillance  
300 for SARS-CoV-2 virus to support public health decisions: Opportunities and challenges. *Current*  
301 *Opinion in Environmental Science & Health* 17:49-71.

302 17. Ahmed W, Tschärke B, Bertsch PM, Bibby K, Bivins A, Choi P, Clarke L, Dwyer J, Edson J, Nguyen  
303 TMH, O'Brien JW, Simpson SL, Sherman P, Thomas KV, Verhagen R, Zaugg J, Mueller JF. 2021.  
304 SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19  
305 transmission in the community: A temporal case study. *Science of The Total Environment*  
306 761:144216.

307 18. Crits-Christoph A, Kantor RS, Olm MR, Whitney ON, Al-Shayeb B, Lou YC, Flamholz A, Kennedy  
308 LC, Greenwald H, Hinkle A, Hetzel J, Spitzer S, Koble J, Tan A, Hyde F, Schroth G, Kuersten S,  
309 Banfield JF, Nelson KL. 2021. Genome Sequencing of Sewage Detects Regionally Prevalent SARS-  
310 CoV-2 Variants. *mBio* 12:e02703-20.

311 19. Fontenele RS, Kraberger S, Hadfield J, Driver EM, Bowes D, Holland LA, Faleye TOC, Adhikari S,  
312 Kumar R, Inchausti R, Holmes WK, Deitrick S, Brown P, Duty D, Smith T, Bhatnagar A, Yeager RA,  
313 Holm RH, Hoogesteijn von Reitzenstein N, Wheeler E, Dixon K, Constantine T, Wilson MA, Lim  
314 ES, Jiang X, Halden RU, Scotch M, Varsani A. 2021. High-throughput sequencing of SARS-CoV-2 in  
315 wastewater provides insights into circulating variants. medRxiv  
316 doi:10.1101/2021.01.22.21250320.

317 20. Martin J, Klapsa D, Wilton T, Zambon M, Bentley E, Bujaki E, Fritzsche M, Mate R, Majumdar M.  
318 2020. Tracking SARS-CoV-2 in Sewage: Evidence of Changes in Virus Variant Predominance  
319 during COVID-19 Pandemic. *Viruses* 12.

320 21. Trujillo M, Cheung K, Gao A, Hoxie I, Kannoly S, Kubota N, San KM, Smyth DS, Dennehy JJ. 2021.  
321 Protocol for Safe, Affordable, and Reproducible Isolation and Quantitation of SARS-CoV-2 RNA  
322 from Wastewater. medRxiv doi:10.1101/2021.02.16.21251787:2021.02.16.21251787.

323 22. Pecson BM, Darby E, Haas CN, Amha YM, Bartolo M, Danielson R, Dearborn Y, Di Giovanni G,  
324 Ferguson C, Fevig S, Gaddis E, Gray D, Lukasik G, Mull B, Olivas L, Olivieri A, Qu Y, Consortium SA-  
325 C-I. 2021. Reproducibility and sensitivity of 36 methods to quantify the SARS-CoV-2 genetic  
326 signal in raw wastewater: findings from an interlaboratory methods evaluation in the U.S.  
327 *Environmental Science: Water Research & Technology* doi:10.1039/D0EW00946F.

- 328 23. Lasek-Nesselquist E, Lapierre P, Schneider E, St. George K, Pata J. 2021. The localized rise of a  
329 B.1.526 variant containing an E484K mutation in New York State. medRxiv  
330 doi:10.1101/2021.02.26.21251868:2021.02.26.21251868.
- 331 24. Annavajhala MK, Mohri H, Zucker JE, Sheng Z, Wang P, Gomez-Simmonds A, Ho DD, Uhlemann  
332 A-C. 2021. A Novel SARS-CoV-2 Variant of Concern, B.1.526, Identified in New York. medRxiv  
333 doi:10.1101/2021.02.23.21252259:2021.02.23.21252259.
- 334 25. CDC.gov. 2021. [https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-](https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html)  
335 [surveillance/variant-info.html](https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html). Accessed March 19.
- 336 26. Deng X, Garcia-Knight MA, Khalid MM, Servellita V, Wang C, Morris MK, Sotomayor-González A,  
337 Glasner DR, Reyes KR, Gliwa AS, Reddy NP, Sanchez San Martin C, Federman S, Cheng J, Balcerek  
338 J, Taylor J, Streithorst JA, Miller S, Kumar GR, Sreekumar B, Chen P-Y, Schulze-Gahmen U, Taha  
339 TY, Hayashi J, Simoneau CR, McMahon S, Lidsky PV, Xiao Y, Hemarajata P, Green NM, Espinosa  
340 A, Kath C, Haw M, Bell J, Hacker JK, Hanson C, Wadford DA, Anaya C, Ferguson D, Lareau LF,  
341 Frankino PA, Shivram H, Wyman SK, Ott M, Andino R, Chiu CY. 2021. Transmission, infectivity,  
342 and antibody neutralization of an emerging SARS-CoV-2 variant in California carrying a L452R  
343 spike protein mutation. medRxiv doi:10.1101/2021.03.07.21252647:2021.03.07.21252647.
- 344 27. Lasek-Nesselquist E, Pata J, Schneider E, George KS. 2021. A tale of three SARS-CoV-2 variants  
345 with independently acquired P681H mutations in New York State. medRxiv  
346 doi:10.1101/2021.03.10.21253285:2021.03.10.21253285.
- 347 28. Posada-Céspedes S, Seifert D, Topolsky I, Jablonski KP, Metzner KJ, Beerenwinkel N. 2021. V-  
348 pipe: a computational pipeline for assessing viral genetic diversity from high-throughput data.  
349 Bioinformatics doi:10.1093/bioinformatics/btab015.

350

351

