1 Partial ORF1ab Gene Target Failure with Omicron BA.2.12.1

- 2 Kyle G. Rodino¹, David R. Peaper², Brendan J. Kelly³, Frederic Bushman⁴, Andrew Marques⁴, Hriju
- 3 Adhikari⁴, Zheng Jin Tu⁵, Rebecca Marrero Rolon⁶, Lars F. Westblade^{6,7}, Daniel A. Green⁸, Gregory J.
- 4 Berry⁸, Fann Wu⁸, Medini K. Annavajhala⁹, Anne-Catrin Uhlemann⁹, Bijal A. Parikh¹⁰, Tracy McMillen¹¹,
- 5 Krupa Jani¹¹, N. Esther Babady^{11,12}, Anne M. Hahn¹³, Robert T. Koch¹³, Nathan D. Grubaugh^{13,14}, Yale
- 6 SARS-CoV-2 Genomic Surveillance Initiative, Daniel D. Rhoads^{5,15,16}
- 7 ¹Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of
- 8 Pennsylvania, Philadelphia, Pennsylvania, USA
- 9 ² Department of Laboratory Medicine, Yale University, New Haven, Connecticut
- ³ Division of Infectious Diseases, Department of Medicine and Department of Biostatistics, Epidemiology,
- and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania,
 USA
- 12 USA
- ⁴ Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia,
- 14 Pennsylvania, USA
- 15 ⁵ Department of Laboratory Medicine, Cleveland Clinic, Cleveland, Ohio
- ⁶ Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, New York
- ⁷ Division of Infectious Diseases, Department of Medicine, Weill Cornell Medicine, New York, New York
- ⁸ Department of Pathology & Cell Biology, Columbia University Irving Medical Center, New York, New
 York
- ⁹ Division of Infectious Diseases, Columbia University Irving Medical Center, New York, New York
- ¹⁰ Department of Pathology and Immunology, Washington University School of Medicine, St. Louis,
 Missouri
- ¹¹ Clinical Microbiology Service, Department of Pathology and Laboratory Medicine, Memorial Sloan
- 24 Kettering Cancer Center, New York, New York
- 25 ¹² Infectious Disease Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New
- 26 York, New York
- ¹³ Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven,
- 28 Connecticut.
- ¹⁴ Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut.
- 30 ¹⁵ Department of Pathology, Cleveland Clinic Lerner College of Medicine, Case Western Reserve
- 31 University, Cleveland, Ohio
- 32 ¹⁶ Infection Biology Program, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio
- 33

34 Corresponding Author:

- 35 Kyle G. Rodino, PhD, D(ABMM)
- 36 Assistant Professor of Pathology and Laboratory Medicine
- 37 3400 Spruce Street
- 38 Hospital of the University of Pennsylvania
- 39 <u>Kyle.rodino@pennmedicine.upenn.edu</u>
- 40 215-662-6652
- 41
- 42 Key Words: Cycle threshold (C₇) value, ORF1ab gene, partial ORF1ab gene target failure (pORF1ab), RT-
- 43 PCR, SARS-CoV-2, whole-genome sequencing

44

45 Abstract

46	Mutations in the viral genome of SARS-CoV-2 can impact the performance of molecular diagnostic
47	assays. In some cases, such as S gene target failure, the impact can serve as a unique indicator of a
48	particular SARS-CoV-2 variant and provide a method for rapid detection. Here we describe partial
49	ORF1ab gene target failure (pOGTF) on the cobas [®] SARS-CoV-2 assays, defined by a ≥2 thermocycles
50	delay in detection of the ORF1ab gene compared to the E gene. We demonstrate that pOGTF is 97%
51	sensitive and 99% specific for SARS-CoV-2 lineage BA.2.12.1, an emerging variant in the United States
52	with spike L452Q and S704L mutations that may impact transmission, infectivity, and/or immune
53	evasion. Increasing rates of pOGTF closely mirrored rates of BA.2.12.1 sequences uploaded to public
54	databases, and, importantly increasing local rates of pOGTF also mirrored increasing overall test
55	positivity. Use of pOGTF as a proxy for BA.2.12.1 provides faster tracking of the variant than whole-
56	genome sequencing and can benefit laboratories without sequencing capabilities.

57

58 Introduction

59 Mutations in the primer/probe binding sites of the SARS-CoV-2 genome can impact oligonucleotide binding and molecular test performance. Throughout the pandemic a number of such mutations have 60 61 been described, resulting in partial or complete PCR target failure (1). Mutations that impair diagnostic 62 detection is a criterion that the United States (US) Center for Disease Control and Prevention (CDC) 63 considers when classifying a novel SARS-CoV-2 lineage as a Variant of Concern (2). Most nucleic acid 64 amplification tests (NAAT) used for in vitro diagnostic (IVD) clinical testing and which have received 65 emergency use authorization (EUA) from the US Food and Drug Administration (FDA) employ multi-66 target assay design. This limits the diagnostic impact of a SARS-CoV-2 mutation that causes single-target

67 failure, as the additional primer/probe target sequences remain unaltered and adequately detect the68 presence of viral nucleic acid.

69 Viral mutations can negatively impact assay performance, but these phenomena have proven useful 70 in certain situations. Whole-genome sequencing (WGS) of SARS-CoV-2 remains inadequately 71 implemented or unavailable in many medical centers across the US due to logistical, cost, and time 72 constraints. Moreover, full genomic characterization of circulating variants often takes multiple weeks 73 from sample collection to data reporting to public health agencies. Specific mutations that yield unique 74 NAAT performance characteristics can provide broader and more rapid assessment of circulating 75 lineages than WGS. For example, the well characterized six base pair deletion in the spike gene resulting 76 in the absence of amino acids H69/V70 (69-70del) results in complete loss of detection of the spike gene 77 target on the Thermo Fisher Scientific TaqPath™ COVID-19 assay (Thermo Fisher Scientific, Waltham, 78 MA), commonly termed S-Gene Target Failure (SGTF). SGTF was a hallmark of the Alpha (B.1.1.7) variant 79 allowing for sequence-free estimation of Alpha's emergence and prevalence in early 2021 (3). Utility in 80 tracking SGTF was observed in late 2021 with the introduction of the Omicron (B.1.1.529.1 [BA.1]) 81 variant, which contained the same characteristic 69-70del. Rapid estimation of Omicron emergence was 82 possible as the preceding dominant variant, Delta, did not harbor the 69-70del and was therefore S-83 Gene Target Positive (SGTP) (4). Finally, in early 2022, SGTF tracking proved useful in monitoring the 84 transition from BA.1 Omicron to BA.2 Omicron (B.1.1.529.2), as the latter lacked the 69-70del and was 85 SGTP, which served as a reliable proxy for identifying BA.2 (5). 86 While the impact of the 69-70del was dramatic, resulting in complete SGTF, other genomic 87 mutations may result in subtler changes including abnormal variance between cycle threshold (C_7) 88 values of multi-target SASR-CoV-2 assays (6-10). Here, we describe partial ORF1ab Gene Target Failure

89 (pOGTF) as detected by abnormal <u>d</u>elta of the C_T values from the <u>E</u> and <u>O</u>RF1ab (dEO) targets of the

90 cobas[®] SARS-CoV-2 assay and cobas[®] SARS-CoV-2 & Influenza A/B assay (Roche Molecular Systems, Inc.;

- 91 Branchburg, NJ) linked to BA.2.12.1, which has rapidly emerged in the US in March-April 2022 and has
- 92 become the dominant SARS-CoV-2 lineage in central New York state (11).

93

94 Materials and Methods

95 Study Population

96 Samples used in this study included upper respiratory specimens (e.g., nasopharyngeal, anterior 97 nares, nasal swabs, and saliva) collected in a transport medium validated by the individual testing 98 laboratory, obtained from patients undergoing standard of care testing for COVID-19 between March 6, 99 2022 and April 16, 2022, a period of time corresponding to MMWR epidemiologic weeks 10 to 15 (12). 100 Samples included in the dEO analysis were tested for the presence of SARS-CoV-2 on cobas® SARS-CoV-2 101 assay or cobas® SARS-CoV-2 & Influenza A/B assays at the originating laboratory. Samples sent for SARS-102 CoV-2 WGS may have been tested on a variety of other FDA EUA-approved SARS-CoV-2 diagnostic 103 platforms, and not a cobas[®] SARS-CoV-2 assay, depending on workflows and testing protocols of the 104 originating laboratory. Laboratories contributing to this study and appropriate IRB determination include 105 the Cleveland Clinic Foundation (CCF) Clinical Microbiology Laboratory (IRB #18-318), the Clinical 106 Microbiology Laboratory of Columbia University Irving Medical Center (CUIMC) (IRB #AAT0123), the 107 Clinical Microbiology and SARS-CoV-2 Molecular Testing Laboratories at the Hospital of the University of 108 Pennsylvania (HUP) (IRB #848605), the Clinical Microbiology Service at Memorial Sloan Kettering (MSK) 109 (IRB #18-491), the Clinical Microbiology Laboratory of Weill Cornell Medical Center (WCMC) (IRB #20-110 03021671), the Barnes-Jewish Hospital Molecular Infectious Disease Laboratory and Washington 111 University in St. Louis (WUSTL) (IRB #20211131), and Clinical Virology Laboratory at Yale-New Haven 112 Hospital (YNHH) and Yale School of Public Health (IRB #2000031374).

113 cobas[®] SARS-CoV-2 assays

114 The cobas[®] SARS-CoV-2 and cobas[®] SARS-CoV-2 & Influenza A/B assays are available on the Roche 115 cobas[®] 6800 and 8800 analyzers (Roche Molecular Systems, Inc.). Both assays perform gualitative 116 detection of SARS-CoV-2 using two genome targets; the ORF1ab gene that is specific to SARS-CoV-2 and 117 the pan-Sarbecovirus envelope (E) gene. Details on the oligonucleotide sequences and the specific 118 genomic regions targeted by the NAAT are not publicly available. MS2 bacteriophage is used as an 119 internal RNA processing control. RNA extraction, reverse transcription, target amplification, and result 120 analysis all occur on the instrument. Positive results are determined by amplification curves that cross a 121 predetermined threshold thus generating C_T values for these loci. These C_T values are available to the 122 laboratory, but are not included in patient reports by any of the performing laboratories. Only 123 qualitative interpretations (SARS-CoV-2 detected, not detected, or presumptive positive) of the results 124 are reported, which is congruent with the EUA instructions for use. Samples do not have to be positive 125 for both gene targets to be called positive by the assay. However, only results positive for both targets 126 were used in this study to allow for C_T value comparison. C_T values used in this study as well as dates of 127 sample collection and/or testing were either directly obtained from instruments or extracted from the 128 laboratory information system depending on the testing laboratory.

129 Identification of samples as dEO outliers

130 Initial recognition of dEO was determined by manual review of C_T values, noticing ≥ 1 or ≥ 2 cycle 131 difference, between the ORF1ab and E C_T values, with ORF1ab C_T value > E gene C_T value, depending on 132 the institution. At HUP we modeled the expected difference in C_{τ} values between two parallel targets as 133 a function of the minimum observed C_{τ} value in the pair, in order to accommodate the observation that 134 the expected difference in C_{τ} values across targets increases at higher C_{τ} values. The expected variation 135 in the C_{τ} value difference was likewise allowed to vary with the minimum observed C_{τ} value in the pair. 136 Bayesian mixed-effect linear regression models, incorporating random effects (slope and intercept) for 137 sequencing platform (i.e., amplicon-target pair), were fit using Stan Hamiltonian Monte Carlo via the

brms package in R (13, 14). Model parameters were established using C_{τ} value data collected across seven platforms from March to June 2021, and the parameterized model was used to monitor for dEO (and other target pair) outliers, with outliers identified as differences beyond the 99% posterior credible interval of the expected C_{τ} value difference for the minimum C_{τ} value observed in the pair. Model code and sample reproducible report are available at: https://github.com/bjklab/SARS-CoV-2 Ct report.

143 SARS

SARS-CoV-2 Whole-Genome Sequencing

144 At CCF, NGS sequencing libraries were prepared with Illumina COVIDSeq Test kit according to the 145 manufacture's recommendation. These libraries were sequenced paired end with read length 151 bases 146 using a NextSeq550 instrument (Illumina, San Diego, CA). Data analysis was performed with an in-house 147 developed bioinformatics pipeline. Sequence reads were mapped to reference genome Wuhan-Hu-1 148 (NC 045512.2) using BWA (version 0.7.15), variant calling was performed with both Freebayes (version 149 1.3.4) and LoFreq (version 2.1.5). Average coverage is x6567 and minimum coverage for mutation is x10. 150 Samples below this minimum coverage at the S gene codon 452 were considered unreliable, and these 151 genotypes were not considered in this study. Manual review of mutations was performed with 152 Integrative Genomics Viewer (IGV) to remove any artifacts. Variant classification was performed with 153 the Pangolin program (https://pangolin.cog-uk.io/ version v4.0.5, lineages version 2022-04-09). Derived 154 genomes with related information were deposited into GISAID database. Samples from (CUIMC) and 155 (WCMC) campuses were sequenced using the Oxford Nanopore (Oxford Nanopore Technologies, 156 Oxford, UK) Midnight protocol targeting 1200bp tiled amplicons across the length of the genome, as 157 previously described (15). Samples were sequenced on an Oxford Nanopore GridION using R9.4.1 flow 158 cells, with negative controls included on each run. Variant calling and consensus genome generation was 159 performed using the Oxford Nextflow ARTIC pipeline. Viral lineage classification, identification of 160 mutations, and phylogenetic analyses were performed using Pangolin v4.0.5 and Nextclade v1.11.0. 161 Genomic data from CUIMC is routinely uploaded to GISAID and to GenBank (under NCBI BioProject

- 162 PRJNA751551). At Yale, sequencing was performed as previously described, but Pangolin v4.0.5 was
- used (16). All other SARS-CoV-2 WGS and lineage assignment was performed as previously described
- 164 (17-19).

165 Data Analysis

- 166 All charts and analyses were generated and performed with GraphPad Prism v8.2.1. Non-linear
- regression of BA.2.12.1 and non-BA.2.12.1 was performed and best fit straight lines using the least
- squares method and *y*-intercepts were compared for differences. Receiver operating characteristic
- 169 (ROC) curve to determine the performance characteristics of different dEO cut-off values was generated
- by classifying samples with lineage information available as either "BA.2.12.1" or "Other". Positive and
- 171 negative predictive values of pOGTF for the detection of BA.2.12.1 were calculated using samples with
- 172 lineage information classified as either "BA.2.12.1" or "Other" where pOGTF was considered a dEO of ≥2
- 173 with an E-gene C_T value \leq 30.

174 Data availability

175 All SARS-CoV-2 viral genomes have been deposited in GISAID and/or NCBI Genbank with accession

176 numbers or viral names listed in Table S1.

177

178 Results

Difference in C_T values for the E gene target compared to the ORF1ab gene target (dEO) from March 6, 2022 to April 16, 2022 were plotted (Figure 1, Table S1). Using an E gene C_T value cutoff of \leq 30 thermocycles to avoid non-specific variation between values near the limit of detection of the assay, an outlier group with abnormal difference between the two C_T values was identified. These data indicate

delayed detection of the ORF1ab target as compared to the E gene target. In total, 428 samples, run on
either the classic or combo SARS-CoV-2 assay, with pOGTF were detected during the study period.

185 Following institution-specific workflows for SARS-CoV-2 surveillance sequencing, a subset of samples 186 from the study period underwent WGS and lineage determination, including 71 samples showing pOGTF 187 (Table S1). Overwhelmingly, pOGTF samples were determined to be BA.2.12.1 (70/71), with the first 188 WGS-confirmed detection of BA.2.12.1 among study participants occurring in the first week of March in 189 the YNHH cohort and remaining study sites also confirming BA.2.12.1 from pOGTF specimens in late-190 March. The emergence of samples with pOGTF coincided with increases in BA.2.12.1, and preliminary 191 evidence suggested that pOGTF may be a marker for BA.2.12.1. To explore this ORF1ab and E gene C_T 192 values were plotted for samples confirmed as BA.2.12.1 or another lineage by sequencing. BA.2.12.1 193 samples also clustered as dEO outliers (Figure 2). Best-fit straight lines demonstrated these to be 194 different populations with significantly different y-intercepts (p < 0.0001). Further analysis of sequencing 195 data demonstrated that BA.2.12.1, a sub-lineage of BA.2, contains five characteristic mutations of 196 interest to this study, compared to the parent strain. These include two missense mutations altering 197 amino acids in the spike gene, g.22917 T>G (S:L452Q) and g.23673 C>T (S:S704L), and three synonymous 198 mutations g.11674 C>T (ORF1ab), g.15009 T>C (ORF1ab), and g.21721 C>T (S). Likely, one of the two 199 ORF1ab synonymous mutations causes reduced efficiency of amplification and/or detection of the 200 target in the cobas[®] SARS-CoV-2 and cobas[®] SARS-CoV-2 & Influenza A/B assays.

We used the subset of samples with lineage information to generate an ROC curve to determine the performance characteristics of different cut-off values for dEO (Figure 3). At a dEO cut-off of 2.1, the sensitivity of was 97.2% while the specificity was 99.9% generated during the ROC analysis. After review of the primary data, there was no difference in performance using a cutoff of 2.0 vs. 2.1, thus dEO \ge 2.0 was selected for simplicity of use in identifying samples with pOGTF (Table 1). When this criterion was applied to our population of samples with lineage information with E gene C_T values \le 30 the positive

and negative predictive values were 98.6% and 99.7%, respectively. Using the thresholds if E gene C₇ values ≤30 and dEO ≥2 to define pOGTF, we reanalyzed our initial cobas SARS-CoV-2 results, finding that the pOGTF group showed a mean dEO of 4.02 (95% CI 3.95 to 4.09) cycles as compared to -0.13 (95% CI -0.14 to -0.11) cycles in the non-pOGTF group (Table 2).

211 To better understand the relationship between pOGTF, the emergence of BA.2.12.1, and overall test positivity, we reviewed publicly available GISAID data to identify the number of all sequences and 212 213 BA.2.12.1 sequences uploaded from North America for the same time period as our study. Over a five-214 week period, the number of sequences submitted was stable, but the proportion of BA.2.12.1 samples 215 increased over time to approximately 13% of all sequences (Figure 4A). Our total testing data using the 216 pOGTF threshold also revealed, similar rise in samples demonstrating pOGTF, with the proportion of 217 positive samples with pOGTF steadily increasing over the study period, peaking in the final week of the 218 study (April 10-16) with a range from 11.1% to 39.8% among the seven study sites (Figure 4B). This rise 219 is continuing. Importantly, the rise in pOGTF at our institutions has also corresponded to an increase in 220 overall test positivity, consistent with concerns of increased transmissibility of the BA.2.12.1 variant.

221 Discussion

222 The role of C_{τ} values in clinical decision-making has been hotly debated throughout the pandemic; 223 however, studies have demonstrated their association with clinical outcomes and thus their potential to 224 inform clinical and logistic decisions (20-30). Nevertheless, multiple sources of pre-analytic and analytic 225 variability, along with the qualitative nature of SARS-CoV-2 assays, can complicate interpretation (31-226 33). However, with enough data points a normal dispersion between C_7 values of multi-target assays can 227 be constructed and monitored for evidence of abnormal differences in C_T values between two targets 228 (Figure 1) (6, 7). Abnormal divergence between C_T values may indicate polymorphisms that impact 229 detection of individual target genes, particularly when the values are well above the limit of detection

(6-10). Outliers showing abnormally large differences between targets or delayed detection of a given
 target relative to other targets may indicate mutations impacting the primer/probe binding sites and
 warrant further characterization by WGS.

233 Tracking unique assay performance characteristics as a marker of a specific variant has 234 demonstrated value through more rapid assessment than WGS, as has been readily apparent with SGTF in the Alpha and Omicron (BA.1 and BA.2) waves of the COVID-19 pandemic. Here we describe the use 235 236 of pOGTF on the cobas® SARS-CoV-2 and cobas® SARS-CoV-2 & Influenza A/B assays to rapidly identify 237 the emergence of BA.2.12.1. pOGTF is a highly reliable proxy for BA.2.12.1, particularly for samples with 238 C_T values \leq 30 thermocycles. This phenomenon is useful for rapid recognition, particularly for 239 laboratories that lack sequencing capabilities. The CDC Nowcast variant proportion projection for the 240 week ending 4/16/22 estimates BA.2.12.1 at 19.0% (13.4-26.0%) nationally, with higher proportions in 241 the Northeast of the US (34). Our data are congruent with these projections, showing highest 242 concentration in the US Northeast with spread into the US Midwest through mid-April, and we 243 demonstrate that pOGTF can be used for real-time BA.2.12.1 tracking. Notably, estimates of the 244 BA.2.12.1 variant fraction using pOGTF are higher than WGS-derived projections. This increased fraction 245 may indicate that BA.2.12.1 is spreading more quickly than has been estimated. 246 The emergence of BA.2.12.1 at our study sites has coincided with increased positivity rates (Figure 247 4), suggesting this variant is at least partially responsible for increased community transmission. 248 Increased transmissibility is likely attributable in part to the spike L452Q mutation present in BA.2.12.1, 249 which is not found in the parent strain but was present in Lambda (C.37). A similar mutation, L452R was 250 observed with Delta (B.1.617.2). L452Q has been implicated as an important driver of human 251 transmission, enhancing infectivity and receptor binding while reducing vaccine-derived immunity (35).

252 For routine diagnostic laboratories, identifying mutations responsible for altering commercial assay 253 performance is hindered by a lack of public information regarding the primer/probe target regions of 254 commercially available FDA EUA assays. As stated during the July 26, 2021 CDC Clinical Laboratory 255 COVID-19 Response Call, "...ideally genomic regions targeted by IVDs would be made publicly available 256 by manufacturers, so prospective investigation of polymorphisms occurring within target regions could 257 be identified" (36). The authors support disclosure of primer/probe target regions for FDA EUA assays to 258 facilitate monitoring for mutations that could impact diagnostic assay performance. As in the case of 259 pOGTF with BA.2.12.1, two putative mutations (g.11674 C>T and g.15009 T>C) have been identified, but 260 the responsible mutation impacting the assay performance has not yet been confirmed. If the former, 261 this supports Wang et al.'s conclusion that cytidines in assay target regions are particularly susceptible 262 to mutations for SARS-CoV-2 (37). Additional challenges for sentinel laboratories in identifying assay-263 impacting mutations are the frequency of the observation and quantity of testing. Sentinel laboratories 264 may struggle to identify an emerging pattern when restricted to local data. It may therefore be easier to 265 identify target failure using larger data sets or by retrospectively retesting samples with mutations 266 identified by WGS, if the genomic regions targeted by the NAAT are publicly available. 267 Currently, mutations impacting primer/probe binding are self-monitored by manufacturers, which 268 may pose a conflict of interest (38). Monitoring therefore relies on publicly generated and openly shared 269 genomic data (GISAID and Genbank). The lag in availability of WGS data may delay detection of new 270 variants that impact assay performance by weeks. Additionally, if a variant emerges in a locale with low 271 sequencing surveillance, recognition may be further delayed. As demonstrated herein, monitoring of assay performance through real-time evaluation of C_T values has utility in SARS-CoV-2 variant 272 273 surveillance, and could inform clinical and logistical decision making. Widespread real-time uploading of 274 SARS-CoV-2 NAAT C_T values to an FDA or manufacturer-sponsored database, similar to the data curation

available for BioFire[®] Syndromic Trends Epidemiology Tool, could permit rapid detection, and thus
response, to emerging variants (39).

277	This study has some limitations. Not all SARS-CoV-2 tests were performed on the cobas [®] SARS-CoV-2
278	assays as multiple NAAT platforms are used at each participating site, and operational workflow may
279	have unintentionally introduced biases (e.g., directing samples from certain patient populations or need
280	for rapid result to particular platforms) to the study set. All analyses were performed on a per-sample
281	and not a per-subject basis, and a single subject's samples could have been included multiple times
282	within the data set. However, the impact of this possibility is diminished by the multi-center nature of
283	this study and large data. While pOGTF is currently specific to BA.2.12.1, the specificity, sensitivity, and
284	predictive values may change over time as SARS-CoV-2 continues to evolve and the lineages prevalent in
285	the population changes. Although pOGTF can be a rapid and useful tool for monitoring BA.2.12.1, WGS
286	remains important in confirming the lineage, assessing for additional mutations, and detecting new
287	variants that decrease the specificity of the association of pOGTF with BA.2.12.1.

288

289 Acknowledgements

290 All authors would like to acknowledge the ASM ClinMicroNet Listserv for providing a platform for 291 communication and collaboration between clinical microbiology laboratory directors, which allowed this 292 group of authors to connect on this shared observation. KGR, BJK, FB, AM acknowledge funding 293 provided by a contract award from the Centers for Disease Control and Prevention (CDC BAA 200-2021-294 10986 and 75D30121C11102/000HCVL1-2021-55232), philanthropic donations to the Penn Center for 295 Research on Coronaviruses and Other Emerging Pathogens, and in part by NIH grant R61/33-HL137063 296 and AI140442 -supplement for SARS-CoV-2. Additional assistance was provided by the Penn Center for 297 AIDS Research (P30-AI045008) and in part with Federal funds from the National Institute of Allergy and

298	Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under
299	Contract No. 75N93021C00015. BJK is also supported by NIH K23 AI121485. They also acknowledge
300	assistance with data collection and analysis by Daniel Danoski, Michael Kosenski, Amee Aghera, Irina
301	Petlakh and the Penn Medicine Data Analytics Center. DAG, GJB, FW, MKA, ACU acknowledge in part
302	NIH funding (U01 DA053949) and the Columbia University Biobank team. DDR and TZ acknowledge
303	Ashley Figula who first recognized the pOGTF phenomenon, which prompted this study; and we thank
304	the Ohio Department of Health for its support of Cleveland Clinic's SARS-CoV-2 genomic
305	surveillance. TM, KJ, NEB acknowledge funding in part through the National Institute of Heath/National
306	Cancer Institute Cancer Center Support (Grant P30 CA008748) and Philanthropic Funds from the Burns
307	Family. NDG acknowledges funding provided by the Centers for Disease Control and Prevention Broad
308	Agency Announcement (75D30120C09570). LFW and RMR acknowledge Kathy Fauntleroy and Selma
309	Salter from New York-Presbyterian/Weill Cornell Medical Center, and Melissa Cushing from Weill Cornell
310	Medicine. We gratefully acknowledge both the originating and submitting laboratories for the sequence
311	data in GISAID EpiCoV on which the SARS-CoV-2 variant prevalence data are partially based.
312	The Yale SARS-CoV-2 Genomic Surveillance Initiative consists of Kendall Billig, Mallery
313	Breban, Chantal Vogels, Kien Pham, Nicholas Chen, Chrispin Chaguza, Irina Tikhonova, Christopher
314	Castaldi, Shrikant Mane, Bony De Kumar, David Ferguson, Nicholas Kerantzas, Marie Landry, Wade
315	Schulz.
316	
317	

318 References

Rhoads DD, Plunkett D, Nakitandwe J, Dempsey A, Tu ZJ, Procop GW, Bosler D, Rubin BP,
 Loeffelholz MJ, Brock JE. 2021. Endemic SARS-CoV-2 Polymorphisms Can Cause a Higher

321		Diagnostic Target Failure Rate than Estimated by Aggregate Global Sequencing Data. J Clin				
322		Microbiol 59:e0091321.				
323	2.	Centers for Disease Control and Prevention. 12-13-2021 2021. SARS-CoV-2 Variant				
324		Classifications and Definitions. <u>https://www.cdc.gov/coronavirus/2019-ncov/variants/variar</u>				
325		classifications.html. Accessed April 18.				
326	3.	Public Health England. 2020. Investigation of novel SARS-CoV-2 variant: Variant of Concern				
327		202012/01.				
328		https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_dat				
329		a/file/959361/Technical Briefing VOC202012-2 Briefing 2.pdf. Accessed April 20.				
330	4.	UK Health Security Agency. 2021. SARS-CoV-2 variants of concern and variants under				
331		investigation in England. Variant of concern: Omicron, VOC21NOV-01 (B.1.1.529): Technical				
332		briefing 30.				
333		https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_dat				
334		a/file/1038404/Technical Briefing 30 pdf Accessed April 18				
335	5.	UK Health Security Agency, 2022, SARS-CoV-2 variants of concern and variants under				
336		investigation in England: Technical briefing 38				
337		https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_dat				
338		a/file/1060337/Technical-Briefing-38-11March2022 pdf Accessed April 20				
330	6	Tahan S. Darikh BA, Droit L. Wallace MA, Burnham CD, Wang D. 2021, SARS-CoV-2 E Gene				
310	0.	Variant Alters Analytical Sensitivity Characteristics of Viral Detection Using a Commercial				
2/1		Povorso Transcription DCP Assay, I Clin Microbiol 50:00007521				
241 242	7	Vanaarschot M. Mann SA. Wahhar IT. Kamm I. Ball SM. Ball I. Hang SN. Nguyan MD. Chan I.V.				
54Z	7.	Phatt KD, Tan M, Datweiler AM, Eccinese A, Wu W, Batson J, Dunorman D, Wadford DA				
545 244		Bildt KD, Tall W, Detwener AW, Espinosa A, Wu W, Batson J, Dynerman D, Watior DA,				
344		Puschinik AS, Neit N, Anyong V, Miller S, Ayscue P, Tato Civi, Paul S, Kistier AL, Dekisi JL, Crawford				
345		ED. 2020. Identification of a Polymorphism in the N Gene of SARS-Cov-2 That Adversely impacts				
346	•	Detection by Reverse Transcription-PCR. J Clin Microbiol 59.				
347	8.	Artesi M, Bontems S, Gobbels P, Franckh M, Maes P, Boreux R, Meex C, Melin P, Hayette MP,				
348		Bours V, Durkin K. 2020. A Recurrent Mutation at Position 26340 of SARS-CoV-2 is Associated				
349		with Failure of the E Gene Quantitative Reverse Transcription-PCR Utilized in a Commercial				
350	-	Dual-Target Diagnostic Assay. J Clin Microbiol 58.				
351	9.	Ziegler K, Steininger P, Ziegler R, Steinmann J, Korn K, Ensser A. 2020. SARS-CoV-2 samples may				
352		escape detection because of a single point mutation in the N gene. Euro Surveill 25.				
353	10.	Hasan MR, Sundararaju S, Manickam C, Mirza F, Al-Hail H, Lorenz S, Tang P. 2021. A Novel Point				
354		Mutation in the N Gene of SARS-CoV-2 May Affect the Detection of the Virus by Reverse				
355		Transcription-Quantitative PCR. J Clin Microbiol 59.				
356	11.	NYSDOH. 2022. New York State Department of Health Announces Emergence of Recently				
357		Identified, Highly Contagious Omicron Subvariants in New York and Urges Continued Vigilance				
358		Against COVID-19. https://www.health.ny.gov/press/releases/2022/2022-04-13_covid-19.htm.				
359		Accessed April 18.				
360	12.	Babady NE, McMillen T, Jani K, Viale A, Robilotti EV, Aslam A, Diver M, Sokoli D, Mason G, Shah				
361		MK, Korenstein D, Kamboj M. 2021. Performance of Severe Acute Respiratory Syndrome				
362		Coronavirus 2 Real-Time RT-PCR Tests on Oral Rinses and Saliva Samples. J Mol Diagn 23:3-9.				
363	13.	Carpenter B GA, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell				
364		A. 2017. Stan: A Probabilistic Programming Language. Journal of Statistical Software 76:1–32.				
365	14.	Burkner P-C. 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of				
366		Statistical Software 80:1-28.				

367 15. Annavajhala MK, Mohri H, Wang P, Nair M, Zucker JE, Sheng Z, Gomez-Simmonds A, Kelley AL, 368 Tagliavia M, Huang Y, Bedford T, Ho DD, Uhlemann AC. 2021. Emergence and expansion of 369 SARS-CoV-2 B.1.526 after identification in New York. Nature 597:703-708. 370 16. Chaguza C, Coppi A, Earnest R, Ferguson D, Kerantzas N, Warner F, Young HP, Breban MI, Billig K, 371 Koch RT, Pham K, Kalinich CC, Ott IM, Fauver JR, Hahn AM, Tikhonova IR, Castaldi C, De Kumar B, 372 Pettker CM, Warren JL, Weinberger DM, Landry ML, Peaper DR, Schulz W, Vogels CBF, Grubaugh 373 ND. 2022. Rapid emergence of SARS-CoV-2 Omicron variant is associated with an infection 374 advantage over Delta in vaccinated persons. Med (N Y) doi:10.1016/j.medj.2022.03.010. 375 17. Everett J, Hokama P, Roche AM, Reddy S, Hwang Y, Kessler L, Glascock A, Li Y, Whelan JN, Weiss 376 SR, Sherrill-Mix S, McCormick K, Whiteside SA, Graham-Wooten J, Khatib LA, Fitzgerald AS, 377 Collman RG, Bushman F. 2021. SARS-CoV-2 Genomic Variation in Space and Time in Hospitalized 378 Patients in Philadelphia. mBio 12. 379 18. Aslam A, Singh J, Robilotti E, Chow K, Bist T, Reidy-Lagunes D, Shah M, Korenstein D, Babady NE, 380 Kamboj M. 2021. Severe Acute Respiratory Syndrome Coronavirus 2 Surveillance and Exposure 381 in the Perioperative Setting With Universal Testing and Personal Protective Equipment Policies. 382 Clin Infect Dis 73:e3013-e3018. 383 19. Chow K, Aslam A, McClure T, Singh J, Burns J, McMillen T, Jani K, Lucca A, Bubb T, Robilotti EV, 384 Babady NE, Kamboj M. 2021. Risk of Healthcare-Associated Transmission of SARS-CoV-2 in 385 Hospitalized Cancer Patients. Clin Infect Dis doi:10.1093/cid/ciab670. 386 20. Binnicker MJ. 2020. Can the Severe Acute Respiratory Syndrome Coronavirus 2 Polymerase 387 Chain Reaction Cycle Threshold Value and Time From Symptom Onset to Testing Predict 388 Infectivity? Clin Infect Dis 71:2667-2668. 389 21. Binnicker MJ. 2021. Can Testing Predict SARS-CoV-2 Infectivity? The Potential for Certain 390 Methods To Be Surrogates for Replication-Competent Virus. J Clin Microbiol 59:e0046921. 391 22. Tom MR, Mina MJ. 2020. To Interpret the SARS-CoV-2 Test, Consider the Cycle Threshold Value. 392 Clin Infect Dis 71:2252-2254. 393 23. Bullard J, Dust K, Funk D, Strong JE, Alexander D, Garnett L, Boodman C, Bello A, Hedley A, 394 Schiffman Z, Doan K, Bastien N, Li Y, Van Caeseele PG, Poliguin G. 2020. Predicting Infectious 395 Severe Acute Respiratory Syndrome Coronavirus 2 From Diagnostic Samples. Clin Infect Dis 396 71:2663-2666. Rhoads DD, Pinsky BA. 2021. The Truth about SARS-CoV-2 Cycle Threshold Values Is Rarely Pure 397 24. 398 and Never Simple. Clin Chem 68:16-18. 399 Magleby R, Westblade LF, Trzebucki A, Simon MS, Rajan M, Park J, Goyal P, Safford MM, Satlin 25. 400 MJ. 2021. Impact of Severe Acute Respiratory Syndrome Coronavirus 2 Viral Load on Risk of 401 Intubation and Mortality Among Hospitalized Patients With Coronavirus Disease 2019. Clin 402 Infect Dis 73:e4197-e4205. 403 Westblade LF, Brar G, Pinheiro LC, Paidoussis D, Rajan M, Martin P, Goyal P, Sepulveda JL, Zhang 26. 404 L, George G, Liu D, Whittier S, Plate M, Small CB, Rand JH, Cushing MM, Walsh TJ, Cooke J, 405 Safford MM, Loda M, Satlin MJ. 2020. SARS-CoV-2 Viral Load Predicts Mortality in Patients with 406 and without Cancer Who Are Hospitalized with COVID-19. Cancer Cell 38:661-671 e2. 407 27. Westblade LF, Magleby R, Trzebucki A, Simon MS, Rajan M, Park J, Goyal P, Safford MM, Satlin 408 MJ. 2021. Reply to Rhoads, et al. Clin Infect Dis 72:e687. 409 Rajyalakshmi B, Samavedam S, Reddy PR, Aluru N. 2021. Prognostic Value of "Cycle Threshold" 28. 410 in Confirmed COVID-19 Patients. Indian J Crit Care Med 25:322-326. 411 29. Rao SN, Manissero D, Steele VR, Pareja J. 2020. A Systematic Review of the Clinical Utility of 412 Cycle Threshold Values in the Context of COVID-19. Infect Dis Ther 9:573-586. 413 30. Miller EH, Zucker J, Castor D, Annavajhala MK, Sepulveda JL, Green DA, Whittier S, Scherer M, 414 Medrano N, Sobieszczyk ME, Yin MT, Kuhn L, Uhlemann AC. 2021. Pretest Symptom Duration

415		and Cycle Threshold Values for Severe Acute Respiratory Syndrome Coronavirus 2 Reverse-
416		Transcription Polymerase Chain Reaction Predict Coronavirus Disease 2019 Mortality. Open
417		Forum Infect Dis 8:ofab003.
418	31.	Rhoads D, Peaper DR, She RC, Nolte FS, Wojewoda CM, Anderson NW, Pritt BS. 2021. College of
419		American Pathologists (CAP) Microbiology Committee Perspective: Caution Must Be Used in
420		Interpreting the Cycle Threshold (Ct) Value. Clin Infect Dis 72:e685-e686.
421	32.	Infectious Diseases Society of America and Association for Molecular Pathology. 2021. IDSA and
422		AMP joint statement on the use of SARS-CoV-2 PCR cycle threshold (Ct) values for clinical
423		decision-making. https://www.idsociety.org/globalassets/idsa/public-health/covid-19/idsa-amp-
424		statement.pdf. Accessed April 19.
425	33.	American Association for Clinical Chemistry. 2021. AACC Recommendation for Reporting SARS-
426		CoV-2 Cycle Threshold (CT) Values AACC.org, on AACC. https://www.aacc.org/science-and-
427		research/covid-19-resources/statements-on-covid-19-testing/aacc-recommendation-for-
428		reporting-sars-cov-2-cycle-threshold-ct-values. Accessed April 19.
429	34.	Centers for Disease Control and Prevention. 2020-03-28 2020. COVID Data Tracker.
430		https://covid.cdc.gov/covid-data-tracker. Accessed April 20.
431	35.	Kimura I, Kosugi Y, Wu J, Zahradnik J, Yamasoba D, Butlertanaka EP, Tanaka YL, Uriu K, Liu Y,
432		Morizako N, Shirakawa K, Kazuma Y, Nomura R, Horisawa Y, Tokunaga K, Ueno T, Takaori-Kondo
433		A, Schreiber G, Arase H, Genotype to Phenotype Japan C, Motozono C, Saito A, Nakagawa S,
434		Sato K. 2022. The SARS-CoV-2 Lambda variant exhibits enhanced infectivity and immune
435		resistance. Cell Rep 38:110218.
436	36.	Rhoads D. Genomic changes causing SARS-CoV-2 detection partial failure, p. In (ed),
437	37.	Wang R, Hozumi Y, Yin C, Wei GW. 2020. Mutations on COVID-19 diagnostic targets. Genomics
438		112:5204-5213.
439	38.	United States Food and Drug Administration, Drug. 2021. Establishing additional Conditions of
440		Authorization for the EUAs of Certain Molecular. Antigen and Serology IVDs related to viral
441		mutations, https://www.fda.gov/media/152406/download, Accessed April 21.
442	39.	Meyers L. Ginocchio CC. Faucett AN. Nolte FS. Gesteland PH. Leber A. Janowiak D. Donovan V.
443		Dien Bard J. Spitzer S. Stellrecht KA. Salimnia H. Selvarangan R. Juretschko S. Daly JA. Wallentine
444		IC. Lindsey K. Moore F. Reed SL. Aguero-Rosenfeld M. Fey PD. Storch GA. Melnick SL. Robinson
445		CC. Meredith IE. Cook CV. Nelson RK. Jones ID. Scarpino SV. Althouse BM. Ririe KM. Malin BA
446		Poritz MA. 2018. Automated Real-Time Collection of Pathogen-Specific Diagnostic Data:
447		Syndromic Infectious Disease Epidemiology, IMIR Public Health Surveill 4:e59
448		



451 Figure 1



453 Plot of ORF1ab and E-gene C^T values generated using the cobas® SARS-CoV-2 or cobas® SARS-CoV-2 & Influenza A/B assays. ORF1ab and E loci C_T values from the indicated institutions for all positive results 454 from March 6, 2022 to April 16, 2022. Vertical line indicates an E-gene C_{T} value of 30 that was used as a 455 456 cut-off for subsequent analyses. Samples with only a single gene detected were not included and are not 457 shown. A total of 3,471 unique samples are included. Abbreviations: Cleveland Clinic Foundation (CCF), 458 Columbia University Irving Medical Center (CUIMC), Hospital of the University of Pennsylvania (HUP), 459 Memorial Sloan Kettering (MSK), Weill-Cornell Medical Center (WCMC), Washington University in St. 460 Louis (WUSTL), and Yale-New Haven Hospital (YNHH). 461





• BA.2.12.1 • Other Lineage

463
 464 Plot of ORF1ab and E-gene C_T values generated using the cobas[®] SARS-CoV-2 or cobas[®] SARS-CoV-2 &

465 Influenza A/B assays for samples with lineage results. ORF1ab and E loci C₇ values from all institutions

for samples with lineage results from March 6, 2022 to April 16, 2022 are shown. All lineages other than
 BA.2.12.1 were consolidated into a single category. A total of 72 BA.2.12.1 and 766 other lineage

BA.2.12.1 were consolidated into a single category. A total of 72 BA.2.12.1 and 766 other lineage
 samples were included. Linear regression was performed, and best fit curves for BA.2.12.1 and all other

469 lineages were had significantly different y-intercepts (p < 0.0001). Only samples with E-gene C_T values

- 470 \leq 30 were used for regression analysis.
- 471

- 472 Figure 3
- 473 Α.



476



481



484 A.



487 Increasing BA.2.12.1 prevalence relates to pOGTF and percent positivity. (A) All North American 488 samples submitted to GISAID from the indicated period of time were extracted on 4/21/22, and the 489 490 percentage of all sequences classified as BA.2.12.1 was calculated. (B) Percent positivity and percent pOGTF for all samples tested by cobas® SARS-CoV-2 or cobas® SARS-CoV-2 & Influenza A/B assays at 491 492 participating institutions for the indicated weeks. For CCF, percent positive data included additional 493 samples tested by an alternative platform. All samples regardless of E-gene C₇ value were included. Percent pOGTF was calculated for all samples with E gene C_T values ≤ 30 with or without a dEO ≥ 2 494 495 cycles.

Table 1. Contingency table and performance analysis using dEO threshold ≥2 for samples with E-gene C_{τ} 498 ≤30 as a marker of BA.2.12.1.

Data analyzed	BA.2.12.1 = Yes	BA.2.12.1 = No	Total
pOGTF = Yes	70	1	71
pOGTF = No	2	765	767
Total	72	766	838

Table Analyzed	pOGTF for BA.2.12.1				
P value and statistical significance					
Test	Fisher's exact test				
P value	<0.0001				
P value summary	* * * *				
One- or two-sided	Two-sided				
Statistically significant (P < 0.05)?	Yes				
Effect size	Value	95% CI			
Sensitivity	0.9722	0.9043 to 0.9951			
Specificity	0.9987	0.9926 to 0.9999			
Positive Predictive Value	0.9859	0.9244 to 0.9993			
Negative Predictive Value	0.9974	0.9905 to 0.9995			
Likelihood Ratio	744.7				

502	Table 2 Descrip	ntive stats for	entire data s	set for same	nles with F	gene (₊ <	30 and dE) >2 (ie	nOGTE)
JUZ	Table 2. Desch	plive stats for	entile uata s	set i Ui saini		gene ut 2		. בב (וב.	pour j.

pOGTF = Yes	pOGTF = No	
428	2213	
2.200	-7.570	
3.600	-0.3000	
4.000	-0.2000	
4.400	0.05000	
10.50	1.900	
8.300	9.470	
95.26%	95.00%	
3.900	-0.2000	
4.040	-0.1900	
4.023	-0.1293	
0.7215	0.3699	
0.03488	0.007863	
3.955	-0.1448	
4.092	-0.1139	
	pOGTF = Yes 428 2.200 3.600 4.000 4.000 4.000 4.000 8.300 8.300 95.26% 3.900 4.040 4.040 4.040 4.023 0.7215 0.03488 0.7215 0.03488	

503

504