

# High-Throughput Adaptable SARS-CoV-2 Screening for Rapid Identification of Dominant and Emerging Regional Variants

Zita Hubler, PhD,<sup>1</sup> Xiao Song, MS,<sup>2</sup> Cameron Norris, MS,<sup>2</sup> Mehul Jani, PhD,<sup>2</sup> David Alouani, PhD,<sup>2</sup> Maureen Atchley,<sup>2</sup> Lisa Stempak, MD,<sup>1,2</sup> Sarah Cherian, MD,<sup>1,2</sup> Christine Schmotzer, MD,<sup>1,2</sup> and Navid Sadri, MD, PhD<sup>1,2,□</sup>

From the <sup>1</sup>Department of Pathology, Case Western Reserve University School of Medicine, Cleveland, OH, USA; and <sup>2</sup>Department of Pathology, University Hospitals Cleveland Medical Center, Cleveland, OH, USA.

## ABSTRACT

**Objectives:** Emerging severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variant strains can be associated with increased transmissibility, more severe disease, and reduced effectiveness of treatments. To improve the availability of regional variant surveillance, we describe a variant genotyping system that is rapid, accurate, adaptable, and able to detect new low-level variants built with existing hospital infrastructure.

**Methods:** We used a tiered high-throughput SARS-CoV-2 screening program to characterize variants in a supraprovincial health system over 76 days. Combining targeted reverse transcription–polymerase chain reaction (RT-PCR) and selective sequencing, we screened SARS-CoV-2 reactive samples from all hospitals within our health care system for genotyping dominant and emerging variants.

**Results:** The median turnaround for genotyping was 2 days using the high-throughput RT-PCR–based screen, allowing us to rapidly characterize the emerging Delta variant. In our population, the Delta variant is associated with a lower cycle threshold value, lower age at infection, and increased vaccine-breakthrough cases. Detection of low-level and potentially emerging variants highlights the utility of a tiered approach.

**Conclusions:** These findings underscore the need for fast, low-cost, high-throughput monitoring of regional viral sequences as the pandemic unfolds and the emergence of SARS-CoV-2 variants increases. Combining RT-PCR–based screening with selective sequencing allows for rapid genotyping of variants and dynamic system improvement.

## INTRODUCTION

Over the past year, new variants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have emerged from all corners of the globe, and some are associated with increased virulence, differing response to treatment, and the ability to evade vaccines.<sup>1</sup> New variants, combined with increased travel and variable vaccination status, can abruptly strain resources.<sup>2,3</sup> There is a need for real-time epidemiologic monitoring of both dominant and low-level SARS-CoV-2 variants on a region-per-region basis in the face of waxing and waning infection rates.<sup>4</sup> Based on the presence of variants with altered transmissibility,

## KEY MESSAGES

- A tiered approach that uses reverse transcription–polymerase chain reaction–based screening to identify dominant variants and sequencing for unique variants maximizes throughput, turnaround time, and information gleaned from each sample.
- In our population, the Delta variant became dominant in less than a month and is associated with lower cycle threshold, lower age at infection, and increased breakthrough cases.
- We identified low-level variants, including the variant of interest B.1.621 and a Delta variant with an E484K mutation in our population using existing laboratory infrastructure.

## KEY WORDS

SARS-CoV-2; Variant identification; Delta variant; Targeted testing; Variant of concern; Variant screening; E484K

*Am J Clin Pathol* June 2022;157:927-935  
[HTTPS://DOI.ORG/10.1093/AJCP/AQAB212](https://doi.org/10.1093/AJCP/AQAB212)

Received: September 4, 2021  
 Accepted: November 16, 2021

Advance publication: January 6, 2022

Corresponding author: Navid Sadri, MD, PhD; [navid.sadri@uhhospitals.org](mailto:navid.sadri@uhhospitals.org).

Funding: Z.H. is supported by NIH grant F30 MH116581-04.

the Centers for Disease Control and Prevention (CDC) and other government organizations have changed public policy and regional guidance.<sup>5</sup> Monitoring of SARS-CoV-2 variants by larger public health entities at the state and federal level relies on smaller sentinel laboratories for variant identification. This identification system needs to be dynamic to quickly identify newly emerging variants, such as Omicron, that may become the dominant strain in a matter of weeks. While it would be ideal, sequencing for variant identification can be cost-prohibitive and not broadly available in many clinical laboratories. By contrast, reverse transcription–polymerase chain reaction (RT-PCR) is available to most clinical laboratories and represents a speedy and cost-effective alternative to genotype common SARS-CoV-2 variants.<sup>6–9</sup> Several versions of RT-PCR–based variant screens detect common genetic alterations in the SARS-CoV-2 sequence to differentiate between variants.<sup>6–9</sup> Yet, RT-PCR–based screens are limited in their availability to detect newly emerging variants that are not identifiable with the existing genomic signatures. This work describes a tiered approach to SARS-CoV-2 variant screening that combines RT-PCR–based screening with the sequencing of select samples to monitor known variants and detect newly emerging variants quickly, accurately, and with flexible infrastructure investment **FIGURE 1A**.

## MATERIALS AND METHODS

### Samples

The study was reviewed by the University Hospitals Institutional Review Board (Cleveland, OH), and ethical approval was waived. SARS-CoV-2 reactive nasal or nasopharyngeal samples identified via routine clinical testing at 1 of 12 regional hospitals between May 21, 2021, and August 4, 2021, were included in the variant screening program. The program was developed to support regional and state variant surveillance efforts. The routine clinical testing platforms varied between regional hospitals and included the following: ID Now COVID-19 (Abbott), Simplexa COVID-19 Direct Kit (DiaSorin), Xpert Xpress SARS-CoV-2 (Cepheid), TaqPath COVID-19 Combo Kit (ThermoFisher Scientific), and Aptima SARS-CoV-2 (Hologic) assays. The Abbott ID NOW COVID-19 and the Cepheid Xpert Xpress SARS-CoV-2 tests were used off-label after validating their use with multiple different viral transport media (phosphate-buffered saline, viral transport media [Ruhof], and Eswab [Copan Diagnostics]). For these assays, use of the assay-specific swabs created logistical problems with ensuring that collection media and site of testing matched. Given that the molecular approaches varied among hospitals, collection sites, and indication, all samples were retested at a centralized facility using a unified RT-PCR approach before being deemed as “positive.”

Deidentified information regarding the patients’ age and vaccination dates within the hospital system was collected from the electronic medical record. For the purposes of this article, vaccine-breakthrough cases were defined as patients who had a reactive SARS-CoV-2 test and had completed their vaccination series within the hospital system more than 2 weeks before the sample collection date. Patients with vaccinations that were not completed within the

hospital system or with incomplete vaccination were not included in the analysis pertaining to vaccination status, such as the analysis in **FIGURE 2D**. Data for the community-wide Cuyahoga County vaccination rates, shown in **Supplemental Figure 1A** (all supplemental materials can be found at *American Journal of Clinical Pathology* online), were downloaded from the Ohio Department of Health COVID-19 Dashboard ([coronavirus.ohio.gov/wps/portal/gov/covid-19/dashboards](https://coronavirus.ohio.gov/wps/portal/gov/covid-19/dashboards)) on August 6, 2021. These community-wide data were not used for the analysis in **FIGURE 2**.

### RNA Extraction and RT-PCR

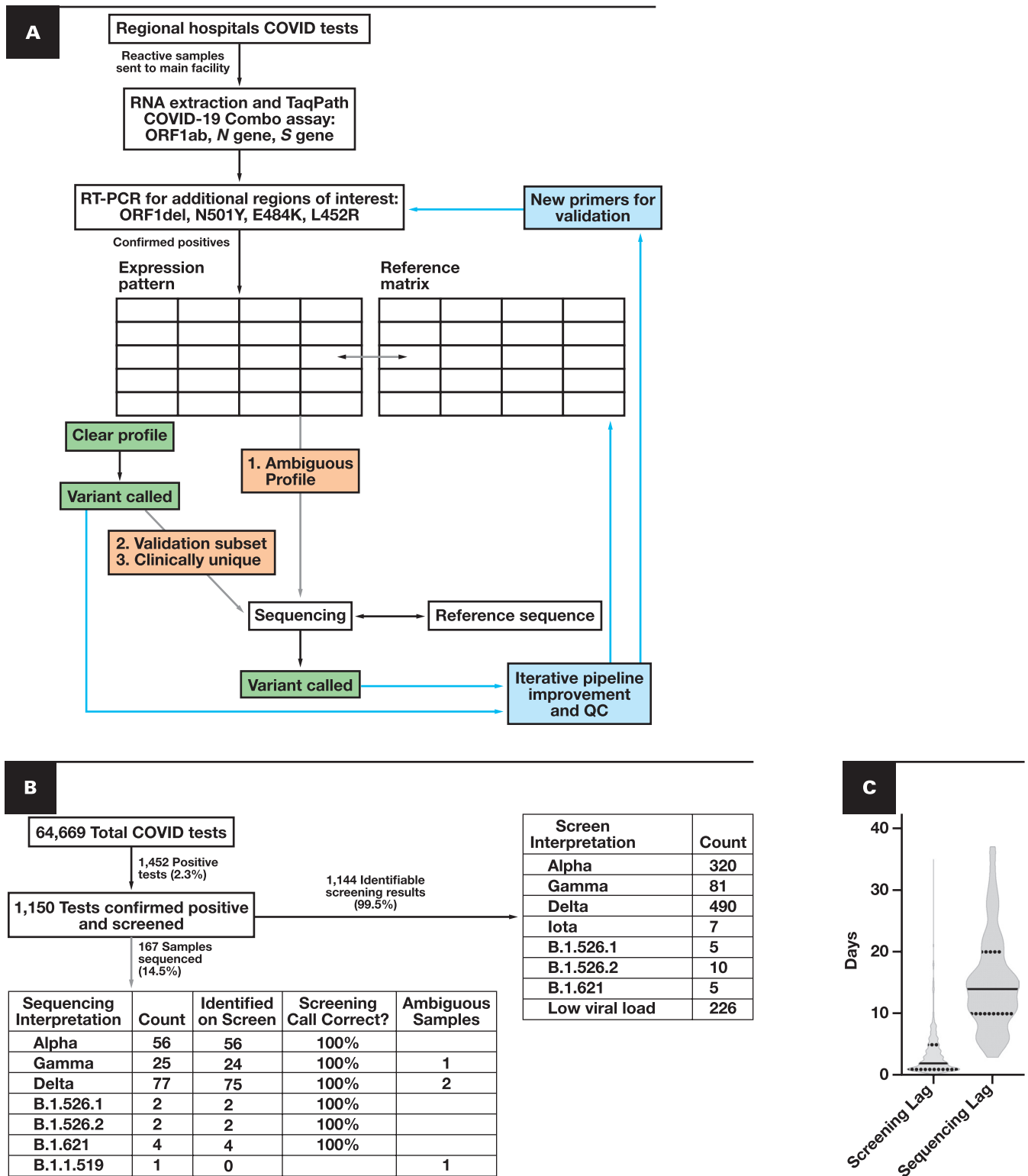
All regional hospitals were instructed to send SARS-CoV-2 reactive samples to a central laboratory for further testing. Samples were retested with the TaqPath COVID-19 Combo Kit assay to enable processing in a high-throughput 96-well plate format. RNA was reextracted and RT-PCRs were performed per the manufacturer’s instructions in a semiautomated fashion using the KingerFisher Flex System (ThermoFisher Scientific) and TaqPath COVID-19 Combo Kit (ThermoFisher Scientific) to assess for sufficient levels of ORF1ab, *N* gene, and *S* gene. The RNA plates were then transported to the sequencing laboratory to perform additional multiplexed RT-PCRs for regions of interest: ORF1del, N501Y, E484K, and L452R. (For primers and more detailed methods, see **Supplemental Table 1**.) Data from all PCR plates were combined and integrated with patient demographic data. A deep learning model was used to identify any instances of errant cycle threshold (Ct) calls.<sup>6</sup> Samples with *N* gene Ct greater than 33 were deemed to have an insufficient viral load for variant calling.

### Variant Preliminary Calling

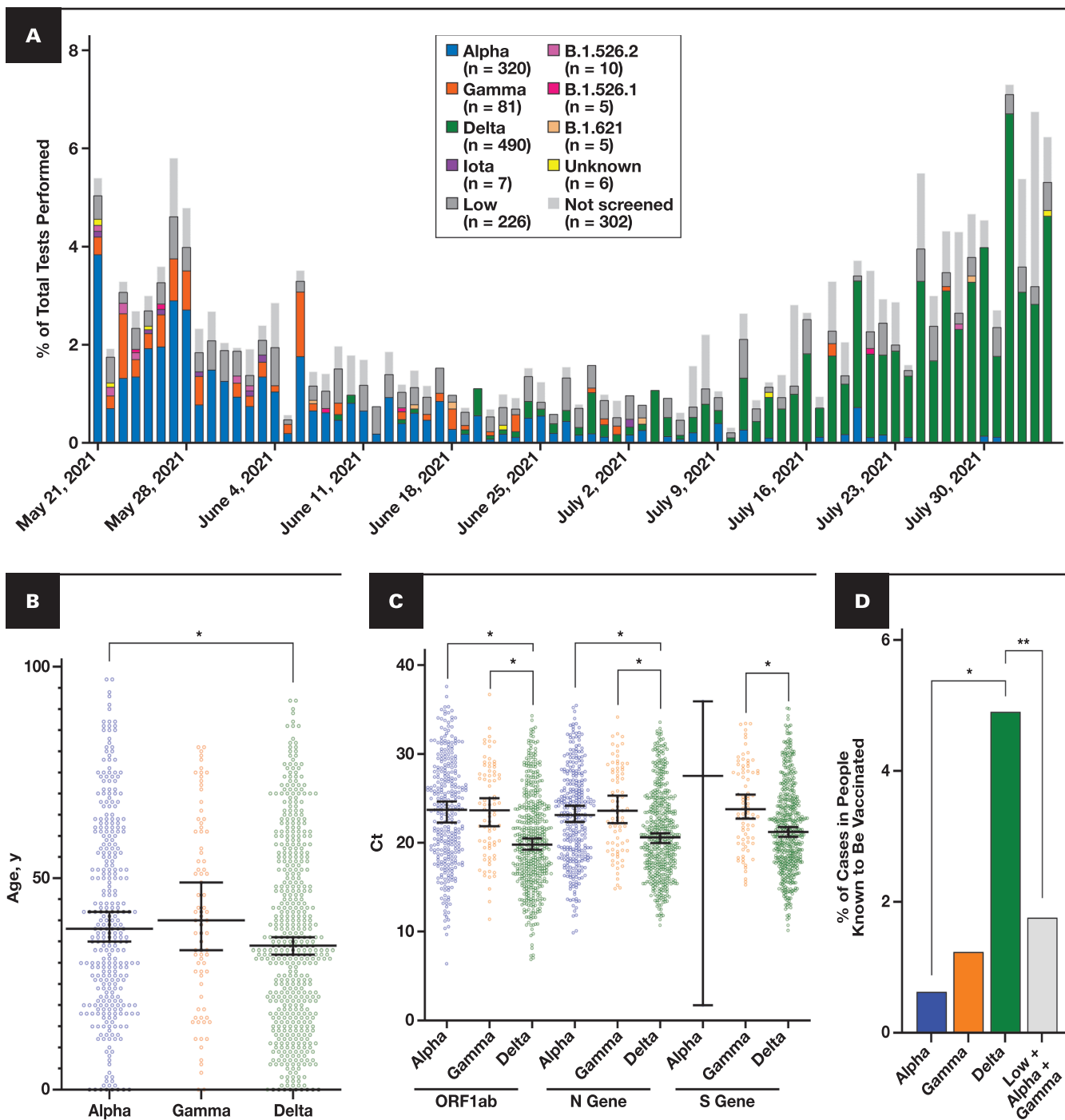
A preliminary probability matrix was calculated using publicly available data that correlated the detection of specific viral genome markers with major World Health Organization (WHO) variants and CDC variants of concern/variants of interest using the primers described. An example of a probability matrix for publicly available data is shown in **Supplemental Table 2**. The preliminary probability matrix was the basis for a ruleset that guides an automated preliminary calling algorithm (**Supplemental Table 3**). The RT-PCR gene profile of each sample and the initial automated variant call are quickly reviewed and finalized **FIGURE 1A**. Samples that have ambiguous expression profiles in comparison to the reference matrix are sequenced for further variant identification **FIGURE 1A**.

### Sequencing

Samples were sequenced for one of three reasons: ambiguous expression profiles relative to the reference matrix on the RT-PCR screen, variant sampling (pipeline validation), or clinically interesting cases (such as vaccine-breakthrough or reinfection cases) **FIGURE 1A**. The number of samples sequenced and sequencing timing are based on the availability of clinical laboratory personnel and resources, as well as the case numbers needed to comprise a sequencing batch. We performed whole-genome sequencing using the SARS-CoV-2 Research Panel (Ion Gene Studio S5 system; ThermoFisher Scientific) as previously



**FIGURE 1** Schematic and flowchart of the dynamic variant calling pipeline. **A**, Schematic of the pipeline. Gray arrows indicate optional pathways, with examples of samples chosen for sequencing. Arrows and boxes coming from variant identification boxes show areas for iterative improvement and quality control of the pipeline. Confirmed positives are samples that have an *N* gene cycle threshold (Ct) of less than 33. **B**, Flowchart of pipeline throughput over 76 days. Data collected from May 21, 2021, through August 4, 2021. Ambiguous samples refer to those samples that the reverse transcription–polymerase chain reaction (RT-PCR) algorithm cannot classify with the current rule set. **C**, Elapsed time between collection, screening, and sequencing for all samples analyzed. Solid line indicates the median; quartiles are indicated with dotted lines. COVID-19, coronavirus disease 2019; QC, quality control.



**FIGURE 2** Tracking of variants and characterization of Delta. **A**, Percentage of each variant per total samples tested per day. Each bar corresponds to 1 day. Variants are color-coded and indicated in the key. The totals indicated in the key are the total count detected via the pipeline over 76 days. “Low” indicates samples with insufficient viral load for screening as measured by a high *N* gene cycle threshold (Ct) greater than 33. “Not screened” samples were not received at the main facility. **B**, Each dot indicates one sample, the y-axis is patient’s age, and x-axis is the variant identified. Median and 95% confidence interval are indicated with the bars. Alpha median age was 38.00, n = 320. Delta median age was 34.00, n = 490. \**P* < .005, Mann-Whitney two-tailed test. **C**, Variant vs Ct for core markers: ORF1ab, *N* gene, and *S* gene. Each dot represents one sample. Median and 95% confidence interval are indicated. \**P* < .0001, one-way analysis of variance with Tukey multiple comparison correction. **D**, Percentage of each variant that is in patients known to be vaccinated relative to all patients identified with that variant. \**P* < .0005, two-tailed Fisher exact test. \*\**P* < .005, two-tailed Fisher exact test.

described.<sup>10</sup> Pangolin (pangolin.cog-uk.io, v.3.1.3, 8/6/2021) and NextClad (clades.nextstrain.org, v.1.5.4) were used for lineage or clade assignment, respectively. Sequencing data were submitted to GenBank (ncbi.nlm.nih.gov/genbank) for public use.

**Data Analysis**

For this study, variant call from the RT-PCR-based Variant Screen and the sequencing results were combined and analyzed using Excel (Microsoft), R (R Foundation), and GraphPad Prism (GraphPad

Software). Statistics were performed using GraphPad Prism. Data for the Delta + E484K samples were downloaded from GISAID ([www.gisaid.org/](http://www.gisaid.org/)) on August 18, 2021.

## RESULTS

### Pipeline Performance

Prior to May 21, 2021, we implemented a ruleset to monitor variants described locally and internationally, as described in the Materials and Methods. Between May 21, 2021, and August 4, 2021, a total of 1,150 positive coronavirus disease 2019 (COVID-19) samples were processed in the RT-PCR–Based Screening Program, representing 2.0% of all COVID-19 tests performed **FIGURE 1B**. In total, 99.5% of the screening results had an identifiable variant expression profile. Of all samples screened, 167 were subsequently sequenced for verification or clarification of variant calling **FIGURE 1B**. Of note, based on our ruleset, the algorithm is not able to differentiate Pangolin lineages of the same variant, such as B.1.617.2 vs AY.3. However, it correctly classified the overall WHO variant as Delta. More granular identification of Pangolin lineages can be accomplished by incorporating novel primers and adjusting the ruleset **FIGURE 1**. For dynamic evaluation of each ruleset iteration, we monitored the validity of the ruleset by sequencing a subset of the samples. We compared the variant calls made by the RT-PCR screen with the variant calls from the sequencing-based clustering. In the samples that were sequenced, the variants called in the screen were correct 100% of the time **FIGURE 1B**. For a subset of samples, the ruleset cannot make a call; these are deemed “ambiguous screening results” and sequenced by default **FIGURE 1B**. These results show that our RT-PCR screening is accurate and high-throughput, a necessary feature given the increasing rate of SARS-CoV-2 infections.

Another essential feature of SARS-CoV-2 variant surveillance is quick turnaround time. We found the median lag time before screening is 2 days after sample collection, with a degree of variability that may be attributable to differing transportation times from various facilities **FIGURE 1C**. By contrast, time to sequencing fluctuates and may be delayed due to practical limitations such as personnel and equipment availability, as well as the number of available samples for a batch. Consequently, the median time between sample collection and sequencing was 14 days, substantially longer than the initial screen **FIGURE 1C**. This discrepancy in median screening vs sequencing time underscores the advantages of a rapid preliminary screen paired with subsequent sequencing.

### Variability in Variant Distribution Over Time

We analyzed the data for the variants called within the past 76 days, as this timeframe represents a relative plateau of vaccination initiation and completion in our county per the Ohio Department of Health, reducing the temporal effects of changing vaccination rates in our community on the analyses performed (**Supplemental Figure 1A**). We noted that in less than a month, the dominant strains shifted from being Alpha and Gamma to Delta **FIGURE 2A**. The transition rate between variants highlights the utility of a rapid variant calling algorithm. Overall, these data parallel the international

trend; after the wave of Alpha swept our county, Delta arrived and was associated with an increase in cases that dominated all other strains in our region.<sup>11,12</sup>

### Impact of the Delta Variant

Given the ongoing concern for Delta’s increasing viral load, infection of younger patients, and vaccine-breakthrough cases, we wanted to characterize the impact of the Delta variant in our population.<sup>11-15</sup> We found a statistically significant decrease in the age of the patients testing positive for Delta compared with Alpha **FIGURE 2B**. Furthermore, we found that Delta samples had a Ct significantly lower than Alpha and Gamma for the three SARS-CoV-2 markers assessed **FIGURE 2C**.<sup>16</sup> This finding may suggest an association between the Delta variant and increased viral load, as Ct has been used as an imperfect surrogate measure for viral load. Last, there was a statistically significant difference in the proportion of patients with each variant who were known to be fully vaccinated within our health care system **FIGURE 2D**.<sup>12</sup> This comparison looks at all vaccinated breakthrough cases within the hospital system and asks what the relative distribution is of each variant. This is expressed as the percentage of samples with each variant over the total within-hospital vaccinated breakthrough cases. This analysis does not include patients who are vaccinated outside our health care system or unvaccinated patients, either in the numerator or denominator; it only includes patients known to be fully vaccinated in our health care system. However, we noted a temporal association between the Alpha variant being dominant and the occurrence of low viral load vaccine-breakthrough events that were not genotyped. This finding suggested that differences in Ct values between Alpha and Delta could be confounding the analysis of vaccine breakthrough (**Supplemental Figure 1B**). Therefore, we assumed that all high Ct breakthrough cases were due to Alpha and Gamma, and despite this, Delta was associated with more vaccinated breakthrough cases **FIGURE 2D**. Together, these results suggest that in our population, the Delta strain correlated with younger patients, lower Ct values, and more vaccine-breakthrough cases.

### Identification of the Novel Delta Variant

One advantage of our process is the ability to triage samples for sequencing that may be more informative. While variants aside from Alpha, Gamma, and Delta represented less than 3.6% of identified cases, these low-level variants can provide clues about the next more virulent or endemic strain<sup>2</sup> **FIGURE 2A**. Unbiased sample sequencing, especially when performed in limited numbers, may miss low-level variants. Our high-throughput RT-PCR screen identified some emerging low-level variants, such as five cases of B.1.621 **FIGURE 2A**. However, novel or unique low-level variants require biased sequencing for further identification.<sup>2</sup> Our identification of a Delta variant with an E484K mutation exemplifies biased sequencing of samples with ambiguous RT-PCR screening results as a mechanism for rapid identification of low-level variants. The sample had an unclear mutational signature on our RT-PCR screen that included an E484K and L452R mutation. Within 2 days, we used sequencing and clustering to elucidate that the sample is a Delta variant based on its Pangolin lineage B.1.617.2 (ambiguity score: 0.9995) and Nextclade clade 21A **FIGURE 3A**. Sequencing confirmed the presence of an E484K





vaccines.<sup>1,13</sup> Information about the presence of newly emerging variants affects public health decision-making, as the variants can overwhelm regional facilities and resources. For example, as the Delta variant arrived in the United States, the recommendations for masking policies from the CDC were modified.<sup>5</sup> Regionally, variant-based surges have led to a call for the preemptive transfer of patients, replenishment of supplies, and increased staff deployment.<sup>4</sup> For low-level variants, if new strains are closely related, they can indicate a common source that can be further investigated.<sup>17</sup> In the United States, larger public health entities rely on hospital laboratories, with various resource levels, for accurate and updated SARS-CoV-2 variant identification. Therefore, hospital laboratories of all resource levels could benefit from rapid and cost-effective variant detection to differentiate between subtly distinct and newly emerging variants. While no one approach is ideal for all facilities, in our hospital system, combining RT-PCR screening and whole-genome sequencing allowed for quick and accurate variant detection that is amenable to optimization.

RT-PCR-based testing is well suited for rapidly identifying known variants with established mutational signatures but is limited in detecting novel variants. Clinical laboratories of all resource levels use RT-PCR-based variant screening due to ease of implementation and cost-efficiency. The widespread uptake of this technology is directly attributable to early work in the area that demonstrated the feasibility of variant identification using a combination of mutational markers, including all or some of the following: L452R, E484K, N501Y, and H69\_V70del.<sup>7-9,18,19</sup> Nevertheless, new variants with unique mutational signatures may not be identifiable with these existing RT-PCR-based screenings, demonstrating the need for further monitoring of low-level or novel variants.

Sequencing is the gold standard for identifying variants and the cornerstone for global monitoring of SARS-CoV-2 viral evolution. In fact, RT-PCR-based monitoring methods rely on global sequencing efforts to stay abreast with the newest viral strains. Resource-rich facilities can use sequencing for variant surveillance on a representative subset of unbiased samples. However, random sampling for sequencing may not be statistically geared to identify low-level variants on a regional basis or rapid enough for public policy purposes. For example, identifying the Delta + E484K variant using unbiased sample selection for sequencing would necessitate sequencing thousands of unmutated Delta variant samples. Turnaround time can be slower, as sequencing runs may be delayed due to demands on laboratory resources when caseloads are high or until batches are filled when caseloads are low, limiting the utility of variant monitoring for real-time outbreak monitoring. To enrich the sequencing data for information, some use targeted sequencing of mutational hotspots or restrict the population to hospitalized patients.<sup>17,20</sup> Nevertheless, monitoring variants exclusively by sequencing is not feasible for most clinical laboratories due to resource limitations, preventing more widespread variant monitoring. Overall, both RT-PCR screening and unbiased sequencing have some capacity to detect novel low-level

variants on a regional basis. Using a tiered approach, we aim to amplify the strengths of each, which may accelerate the identification of emerging clinically important strains.

We implemented a pipeline with a RT-PCR-based screen for the most common variants, with a median turnaround time of 2 days for 99.5% of samples, and subsequent sequencing of select samples. Examples of sequenced samples include ambiguous RT-PCR expression profiles (such as Delta with an E484K mutation), clinically interesting cases (such as vaccine-breakthrough or reinfection cases), or quality control samples. Depending on available resources, sequencing can be performed in-house or outsourced to other facilities, such as companies or other laboratories. In short, this tiered approach aims to match the strengths of each technology to the most informative samples.

The merits of this system are exemplified by our timely detection of rare variants while maintaining high-throughput variant identification of the dominant strains. Due to the high-throughput nature of the RT-PCR screen, we identified the emerging Delta variant and, within the first few weeks, collected sufficient samples to perform preliminary characterization of it in our population. We found the Delta variant is associated with younger patient age, lower Ct values, and increased vaccine breakthrough. As a testament to the utility of sequencing, we were able to detect a rare and emerging variant of Delta with an E484K mutation in 2 days, which had not yet been described in our population and only a handful of times previously in the United States. Therefore, our approach demonstrates the strengths of both RT-PCR-based screening and sequencing and outlines the practical implementation of these tests using existing clinical laboratory infrastructure for the purposes of broader SARS-CoV-2 variant monitoring.

In addition to the benefits gleaned from the combined advantages of the RT-PCR and sequencing approaches, one unique advantage of the tiered approach is its capacity for dynamic quality control and optimization. As the virus evolves, so too must the tests we use to detect it. In this study, we preemptively implemented RT-PCR primers and a ruleset for the detection of variants that were being monitored globally, including the Delta variant (Supplemental Table 3). By pairing RT-PCR with sequencing, we evaluated the efficacy of the approach dynamically once Delta arrived in our community and showed that the RT-PCR screen was correct in variant calling. Despite this, a Delta E484K was not identified based on our screen due to an ambiguous profile. Going forward, the next iteration of the ruleset could include Delta with an E484K for rapid detection and monitoring. Therefore, successful implementation of iterative improvement and dynamic evaluation is shown through the detection of the Delta variant. Furthermore, identification of a Delta E484K with sequencing, despite being unidentifiable in the initial screen, demonstrates the need for additional iterations and continuous sequencing-based quality control. Given that the SARS-CoV-2 genome is accumulating mutations and new variants are evolving, stagnant screening systems can become obsolete.<sup>21-23</sup> As an example, it is expected that the proposed screen approach would effectively detect the most recent variant of concern, Omicron,

based on current understanding of its genetic sequence. A tiered approach allows for continuous quality control and improvement to match the pace of variant evolution.

As with all approaches, there are some possible drawbacks to the tiered system. The pipeline will not catch all mutations or new variants; instead, it is designed to be adaptable and broadly implementable. First, if a variant is undetectable with a specific preliminary COVID-19 assay, that sample would not be further processed in this pipeline for variant identification. The rate of false negatives is directly dependent on the reliability of the preliminary COVID-19 screen done at the regional hospital, which can vary.<sup>24</sup> Furthermore, new genetic variants may have specific mutations that allow it to evade detection by a particular molecular assay or the RT-PCR screen. An alternative source of variants is if a new strain or substrain has novel mutations that are functionally relevant but do not lead to altered classification by the RT-PCR-based screen. In these circumstances, if the new mutation-containing variant becomes more prevalent, it should be detected based on the quality control sequencing done and flag review of the current ruleset. Another drawback is that if biased sequencing is used broadly, it may lead to misrepresentation of low-level variants in sequencing repositories, such as GISAID, that rely purely on regional sequencing results to monitor the global levels of variants. If biased sequencing were broadly implemented, the repositories could address this concern by having submitters indicate if bias or unbiased sampling was performed. As with all variant identification approaches, these limitations need to be weighed against the need for throughput, speed, decreased resource usage, and complexity of data.

The tiered approach is not intended to replace sequencing; it is an alternative approach for facilities to optimize their existing resources. Given the variability of hospital laboratory resources, the suitability of an approach is context dependent. For example, for high-resource laboratories or low-volume laboratories, sequencing may be most appropriate, as it is the gold standard. By contrast, RT-PCR-based approaches are more affordable than either sequencing or our tiered system and may be more suitable for very high-volume or low-resource laboratories. The main advantage of the tiered workflow is that it maintains the benefits of both sequencing-based (such as detecting novel mutants) and RT-PCR-based (such as speed and throughput) variant identification. As demonstrated by our monitoring of the Delta variant's arrival and the detection of a unique Delta + E484K sample, the tiered approach can optimize clinical laboratory resources to accommodate the need for high-throughput rapid variant identification for epidemiologic surveillance of SARS-CoV-2.

**Acknowledgments:** We are grateful for the technical, medical, and administrative staff at University Hospitals Health System for their unwavering support and resilience of our SARS-CoV-2 testing and variant identification programs. All data relevant to the study are included in the article or uploaded as supplementary information. Sequencing data have been submitted to GenBank.

## REFERENCES

- Centers for Disease Control and Prevention (CDC). SARS-CoV-2 variant classifications and definitions. <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html>. Accessed August 18, 2021.
- Snell LB, Cliff PR, Charalampous T, et al. Rapid genome sequencing in hospitals to identify potential vaccine-escape SARS-CoV-2 variants. *Lancet Infect Dis*. 2021;21:1351-1352.
- Xia F, Yang X, Cheke RA, et al. Quantifying competitive advantages of mutant strains in a population involving importation and mass vaccination rollout. *Infect Dis Model*. 2021;6:988-996.
- Kadri SS, Simpson SQ. Potential implications of SARS-CoV-2 Delta variant surges for rural areas and hospitals. *JAMA*. 2021; 326:1003-1004.
- Centers for Disease Control and Prevention (CDC). Delta variant: what we know about the science. <https://www.cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html>. Accessed August 18, 2021.
- Alouani DJ, Rajapaksha RRP, Jani M, et al. Specificity of SARS-CoV-2 real-time PCR improved by deep learning analysis [published online May 19, 2021]. *J Clin Microbiol*.
- Korukluoglu G, Kolukirik M, Bayrakdar F, et al. 40 Minutes RT-qPCR assay for screening spike N501Y and HV69-70del mutations [posted online January 26, 2021]. *bioRxiv*.
- Vogels CBF, Breban MI, Ott IM, et al; Brazil-UK CADDE Genomic Network; Network for Genomic Surveillance in South Africa. Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. *PLoS Biol*. 2021;19:e3001236.
- Wang H, Miller JA, Verghese M, et al. Multiplex SARS-CoV-2 genotyping reverse transcriptase PCR for population-level variant screening and epidemiologic surveillance. *J Clin Microbiol*. 2021;59:e0085921.
- Sekulic M, Harper H, Nezami BG, et al. Molecular detection of SARS-CoV-2 infection in FFPE samples and histopathologic findings in fatal SARS-CoV-2 cases. *Am J Clin Pathol*. 2020;154:190-200.
- Mishra S, Mindermann S, Sharma M, et al; COVID-19 Genomics UK (COG-UK) Consortium. Changing composition of SARS-CoV-2 lineages and rise of Delta variant in England. *Eclinicalmedicine*. 2021; 39:101064.
- Christensen PA, Olsen RJ, Long SW, et al. Delta variants of SARS-CoV-2 cause significantly increased vaccine breakthrough COVID-19 cases in Houston, Texas [posted online July 22, 2021]. *medRxiv*.
- Kirola L. Genetic emergence of B.1.617.2 in COVID-19. *New Microbes New Infect*. 2021;43:100929.
- Lopez Bernal J, Andrews N, Gower C, et al. Effectiveness of Covid-19 vaccines against the B.1.617.2 (delta) variant. *N Engl J Med*. 2021;385:585-594.
- Li B, Deng A, Li K, et al. Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 delta variant [posted online July 23, 2021]. *medRxiv*.
- Alaimo JT, Besse A, Alston CL, et al. Loss-of-function mutations in ISCA2 disrupt 4Fe-4S cluster machinery and cause a fatal leukodystrophy with hyperglycemia and mtDNA depletion. *Hum Mutat*. 2018;39:537-549.
- Page AJ, Mather AE, Le-Viet T, et al. Large-scale sequencing of SARS-CoV-2 genomes from one region allows detailed epidemiology and enables local outbreak management. *Microb Genom*. 2021;7:000589.
- Zhao Y, Lee A, Composto K, et al. A novel diagnostic test to screen SARS-CoV-2 variants containing E484K and N501Y mutations. *Emerg Microbes Infect*. 2021;10:994-997.
- Wang H, Jean S, Eltringham R, et al. Mutation-specific SARS-CoV-2 PCR screen: rapid and accurate detection of variants of concern and the identification of a newly emerging variant with spike L452R mutation. *J Clin Microbiol*. 2021;59:e0092621.



20. Jørgensen TS, Blin K, Kuntke F, et al. A rapid, cost efficient and simple method to identify current SARS-CoV-2 variants of concern by Sanger sequencing part of the spike protein gene [posted online March 29, 2021]. *medRxiv*.
21. Bandoy DJDR, Weimer BC. Analysis of SARS-CoV-2 genomic epidemiology reveals disease transmission coupled to variant emergence and allelic variation. *Sci Rep*. 2021;11:7380.
22. Jonsson H, Magnusson OT, Melsted P, et al. Molecular benchmarks of a SARS-CoV-2 epidemic. *Nat Commun*. 2021;12:3633.
23. Sjaarda CP, Guthrie JL, Mubareka S, et al. Temporal dynamics and evolution of SARS-CoV-2 demonstrate the necessity of ongoing viral genome sequencing in Ontario, Canada. *mSphere*. 2021;6:e00011-21.
24. US Food and Drug Administration. SARS-CoV-2 reference panel comparative data. <https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/sars-cov-2-reference-panel-comparative-data>. Accessed October 3, 2021.