

# Genotyping SARS-CoV-2 Variants Using Ratiometric Nucleic Acid Barcode Panels

Hannah N. Kozlowski,<sup>✉</sup> Ayden Malekjahani,<sup>✉</sup> Vanessa Y. C. Li, Ayokunle A. Lekuti, Stephen Perusini, Natalie G. Bell, Veronique Voisin, Delaram Pouyabahr, Shraddha Pai, Gary D. Bader, Samira Mubareka, Jonathan B. Gubbay, and Warren C. W. Chan\*



Cite This: <https://doi.org/10.1021/acs.analchem.2c04630>



Read Online

ACCESS |



Metrics & More

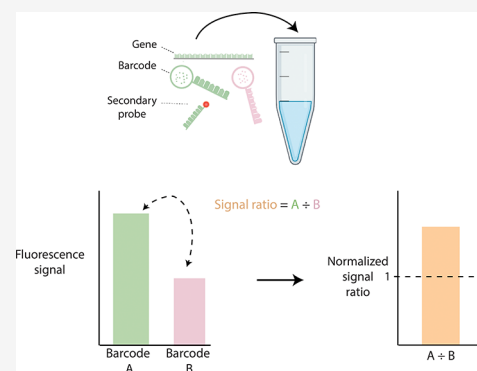


Article Recommendations



Supporting Information

**ABSTRACT:** Designing diagnostic assays to genotype rapidly mutating viruses remains a challenge despite the overall improvements in nucleic acid detection technologies. RT-PCR and next-generation sequencing are unsuitable for genotyping during outbreaks or in point-of-care detection due to their infrastructure requirements and longer turnaround times. We developed a quantum dot barcode multiplexing system to genotype mutated viruses. We designed multiple quantum dot barcodes to target conserved, wildtype, and mutated regions of SARS-CoV-2. We calculated ratios of the signal output from different barcodes that enabled SARS-CoV-2 detection and identified SARS-CoV-2 variant strains from a sample. We detected different sequence types, including conserved genes, nucleotide deletions, and single nucleotide substitutions. Our system detected SARS-CoV-2 patient specimens with 98% sensitivity and 94% specificity across 91 patient samples. Further, we leveraged our barcoding and ratio system to track the emergence of the N501Y SARS-CoV-2 mutation from December 2020 to May 2021 and demonstrated that the more transmissible N501Y mutation started to dominate infections by April 2021. Our barcoding and signal ratio approach can genotype viruses and track the emergence of viral mutations in a single diagnostic test. This technology can be extended to tracking other viruses. Combined with smartphone detection technologies, this assay can be adapted for point-of-care tracking of viral mutations in real time.



Rapidly mutating viruses have emerged, circulated, and even disappeared worldwide for hundreds of years. The mutational patterns of a virus' genome define its genotype, which has implications in disease development, immune responses, therapeutic efficacies, and host outcomes.<sup>1–5</sup> The process of identifying the mutations a virus has collected is referred to as genotyping. Genotyping is used to monitor rapidly mutating viruses in the population.<sup>6</sup> Technologies for viral genotyping include next-generation sequencing and nucleic acid assays such as polymerase chain reaction (PCR).<sup>7,8</sup> Next-generation sequencing comprehensively reads entire pathogen genomes in a single run but requires large infrastructure and has long turnaround times.<sup>9–13</sup> Nucleic acid molecular assays are less resource-intensive genotyping tools than sequencing,<sup>14–16</sup> but are fundamentally limited to a single target per reaction.<sup>7</sup> Nucleic acid genotyping assays use a series of single target tests that first confirm the identity of a virus, followed by mutational analysis to genotype a viral variant. However, genotyping sequences with a single base pair mutation using only a single nucleic acid probe is difficult.<sup>17,18</sup> A single base pair change does not have a large effect on probe stability and will produce a comparable level of optical readout signal to a nonmutated sequence.<sup>18</sup> Researchers attempt to address this problem by

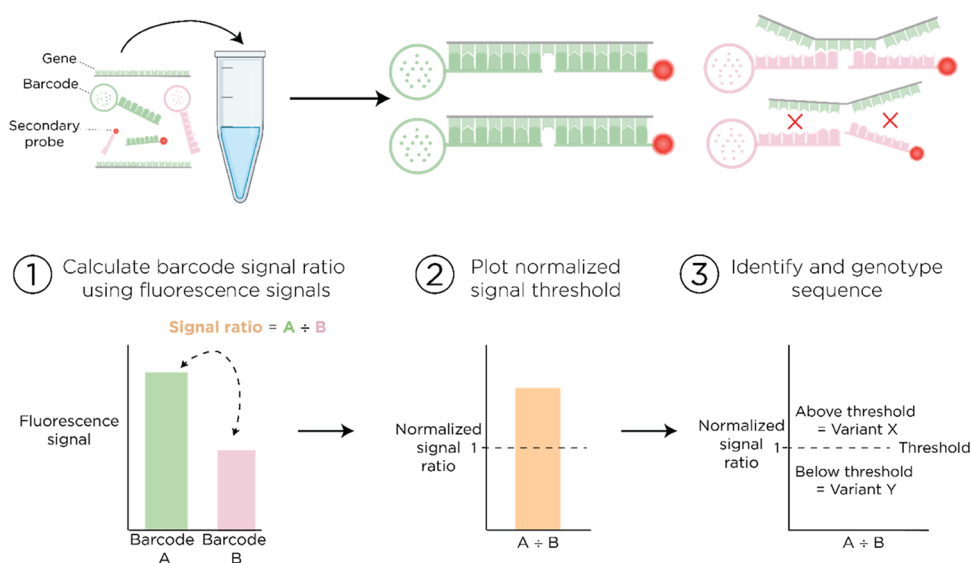
designing probes using modified nucleotides that improve base pair specificity<sup>18,19</sup> and changing assay conditions to favor the binding of one sequence over another, but these approaches lead to complex probe design and optimization processes.<sup>20,21</sup> We developed a simplified and multiplex quantum dot barcoding strategy for genotyping viruses. Scheme 1 describes the barcode assay and strategy.

## EXPERIMENTAL SECTION

**Quantum Dot Bead Synthesis and Conjugation.** The quantum dot (QD) barcodes were synthesized as per our established concentration-controlled flow focusing method.<sup>22</sup> Briefly, four separate QD barcodes were prepared by mixing the QDs, iron(II) oxide nanoparticles (10mg/mL, NN Crystal), and polymer solution (poly(styrene-co-maleic anhydride), 4% w/v, Thermo Fisher) together in different ratios (Table S1). SARS-

Received: October 19, 2022

Accepted: February 6, 2023

Scheme 1. Genotyping Workflow Using Barcode Assay and Signal Ratios<sup>a</sup>

<sup>a</sup>A gene is mixed with two barcodes: barcode A (green) and barcode B (pink). The green gene preferentially binds to perfectly complementary barcode A and partially binds to partially complementary barcode B. The barcode signals are measured using flow cytometry. (1) Signal ratios are calculated by dividing the fluorescence signals of barcode A by barcode B. (2) A signal threshold is calculated using the negative conditions. The threshold value (dotted line) defines the identity of the sequence. (3) A sequence is identified based on whether it is below or above the threshold.

CoV-2 identification probes were designed by aligning whole SARS-CoV-2 genomes. Each QD barcode was conjugated to a specific oligonucleotide identification probe (Table S2) using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) cross-linker (Sigma) chemistry. Each QD barcode was conjugated to the corresponding identification sequence for 4 h at room temperature in a rotator and washed three times using 0.05% (w/v) Tween-20 to eliminate any unbound identification probe.

**Clinical Sample Collection, Extraction, and Amplification.** Deidentified nasopharyngeal swab clinical samples were obtained from local health units (Toronto, Ontario) and stored at  $-80^{\circ}\text{C}$ . RNA from nasopharyngeal swabs was extracted using EasyMag<sup>TM</sup> (bioMerieux, France) magnetic extraction. Samples were reverse-transcribed using the ProtoScript First Strand cDNA Synthesis kit (New England BioLabs), then isothermally amplified with TwistAmp Basic RPA kit (TwistDx), and purified using the GeneJET PCR Purification Kit (Thermo Fisher, Canada). Reverse transcription and purification were done using manufacturer's protocols. For duplex RPA, 59  $\mu\text{L}$  of rehydration buffer, 4.8  $\mu\text{L}$  of each E gene primer, 3.2  $\mu\text{L}$  of each RdRP gene primer (10  $\mu\text{M}$ ), 10  $\mu\text{L}$  of 280 mM MgOAc, and two RPA pellets were mixed with the target. The volumes and materials above were halved for a singleplex RPA for the spike gene amplification. Reactions were then incubated at  $40^{\circ}\text{C}$  for 40 min before nucleic acid purification. Primers were designed using whole genome alignments or modified from the literature (Table S2).<sup>23,24</sup> All clinical samples were processed in duplicate.

**QD Barcode Probe Design.** Two probes were used to detect each target in a sandwich assay as described by Kim et al.<sup>25</sup> The two probes were the reporter probe and the identification probe. Reporter probes were 3'-Cy5-5'-functionalized oligos that confirmed that the sequence of interest was present. The identification probes for the E, RdRP, and spike genes were designed to bind a region within each amplicon. Identification sequences for spike gene were designed with 0–2 mutations in the center as it decreases probe binding.<sup>4</sup> Three

different identification sequences were tested, compared to existing sequences, and the best-performing probes with minimal cross-reactivity with other probes were selected (Table S2).

**Bioinformatics Analysis.** The number of mutations between probes and viral sequences was tracked by analyzing SARS-CoV-2 genomes from NCBI and GISAID databases. Genomes from GISAID and NCBI were downloaded each month between April 2020 and March 2021. Genome sequences were selected based on the following criteria: complete, high coverage, excluding low coverage and from humans. Sequences of lower quality were excluded from analysis including sequences with >550 Ns, >10 gaps and a length >30 kb. Analysis was performed using shell-scripting and R. Multiple sequence alignment (MSA) was performed using MAFFT version 7 (<https://mafft.cbrc.jp/alignment/software/>). We aligned all genomes and as well as the probe sequences using default parameters. MAFFT has an iterative alignment algorithm and is useful for sequences containing large gaps. MSAs were visually evaluated using UniProt UGENE (<http://ugene.net/>). This visualization permitted an assessment of the probe-genome alignment region to identify systematic gaps in the reference genome MSA. trimAl was used to clean the MSA and remove regions with all gaps. MSA results were then imported into R and the Biostrings\_2.62.0 package was used to count the number of mismatches of each probe against the reference genome sequence "Wuhan/WIV04/2019" and logos were generated using the ggseqlogo package. Software created for this analysis is available on GitHub at <https://github.com/BaderLab/Chan-Covid19>.

**Detection of RPA Amplicons with QD Barcodes Assay.** RPA amplicons (2  $\mu\text{L}$ ) were denatured at  $95^{\circ}\text{C}$  for 5 min. For the singleplex barcode assays, amplicons were mixed with 10,000 microbeads, 28 pmol of Cy5-functionalized reporter oligonucleotides, 10  $\mu\text{L}$  of hybridization buffer (4 $\times$  SSC, 0.05% SDS), and water or target to make a total reaction volume of 20  $\mu\text{L}$ . For a multiplex assay, 2500 of each type of barcode, 5 pmol of each

**Table 1. Barcoding Panels and Targets**

barcoding system	barcode 1	barcode 2	barcode 3	barcode 4
conserved genes	SARS-CoV-2 E gene	SARS-CoV-2 RdRP gene	noncomplementary gene #1	noncomplementary gene #2
H69/S70 deletion	SARS-CoV-2 E gene	noncomplementary gene	SARS-CoV-2 spike gene H69/S70 wildtype	SARS-CoV-2 spike gene H69/S70 deletion
N501Y substitution	SARS-CoV-2 E gene	noncomplementary gene	SARS-CoV-2 spike gene N501 wildtype	SARS-CoV-2 spike gene N501Y mutant

reporter, 5  $\mu$ L of hybridization buffer (8x SSC, 0.2% SDS), and water or target were combined. Each reaction was rotated at 37 °C for 25 min and washed using magnetic separation (Life Technologies) for 10 min in the presence of wash buffer (0.5x SSC, 0.1% SDS). Samples were resuspended in 250  $\mu$ L of phosphate-buffered saline–tween (PBS-T) buffer and analyzed on flow cytometry (Figures S1 and S2, (LSR Fortessa, BD)).

**Flow Cytometry Data Analysis.** Data were analyzed as previously described by Kim et al.<sup>25</sup> Briefly, all microbead populations were gated in FlowJo software (BD) using forward scatter and side scatter. Different quantum dot barcodes were gated using the violet 605/20 and violet 525/50 filters. The red 670/14 filter was used to measure the Cy5 secondary probe signal and determine whether target was present on each bead. The signal from the red 670/14 filter was compared between barcodes to calculate ratios.

**Determining Threshold Values.** Table 1 describes the barcode panel designs for each experiment. Thresholds for the detection of conserved genes in clinical sample experiments were made as follows.

The E, RdRP, and noncomplementary gene #1 were incubated in water. For the E gene signal ratio, the signal from the E gene barcode was divided by the signal from the noncomplementary gene #1 to create the E gene signal ratio. The replicates of the E gene signal ratios for the water condition were averaged. The threshold was drawn by taking three standard deviations above the average E gene signal ratios for the water condition. The threshold for the RdRP detection was drawn similarly. Briefly, the RdRP signal ratio was calculated by dividing the RdRP barcode signal over the signal from the noncomplementary #1 barcode for the water condition. The RdRP signal ratios for the replicates were averaged, and the threshold was calculated by taking three standard deviations above this value.

Thresholds for the H69/S70 four-barcode panel were made as follows. The E gene threshold was calculated the same way as the E gene threshold in the conserved gene experiment. The H69/S70 spike gene detection threshold was calculated using the spike gene signal ratio. The spike gene signal ratios were calculated using SARS-CoV-2-positive spike mutant samples. These samples were 1000, 100, and 10 nM samples that contained both E gene and H69/S70 deletion spike gene. For these three concentrations, the wildtype barcode signal was divided by the mutant barcode signal for each replicate. The average of the spike gene signal ratios for three concentrations was taken. We drew the threshold by taking three standard deviations for the three concentrations and adding them to the calculated average signal ratio.

Thresholds for the N501Y barcode assay for detecting synthetic samples were calculated as follows. The E gene threshold was calculated the same way as in the conserved gene experiment. The N501Y spike gene detection threshold was calculated using the N501Y spike gene signal ratio. The N501Y

signal ratios were calculated using the SARS-CoV-2-positive samples that contained the N501Y mutant. These samples included  $10^6$ – $10^2$  copies of both amplified E and spike N501Y mutant gene. For these three concentrations, the wildtype barcode signal was divided by the mutant barcode signal for each of the replicates. The average of the spike gene signal ratios for five concentrations was taken. We drew the threshold by adding three standard deviations of the spike gene signal ratio to the calculated average signal ratio.

## RESULTS AND DISCUSSION

### Our Barcoding and Viral Mutation Detection Strategy.

A central concept in nucleic acid thermodynamics is that the large number of hydrogen bonds between complementary single-stranded sequences makes it energetically favorable to bind to one another.<sup>23–28</sup> Partly complementary single-stranded sequences have a lower number of hydrogen bonds between them and form fewer double-stranded products.<sup>40,41</sup> We hypothesize that we can identify and genotype viral variants by comparing the amount of double-stranded products formed when perfectly complementary and mismatched probe sequences bind to a variant.

Our strategy uses quantum dot barcodes, which are polymeric microbeads embedded with different fluorescence-emitting quantum dots. We designed a panel of barcodes using multiple microbeads with unique fluorescence emission profiles to detect multiple targets simultaneously. Each uniquely emitting optical bead is conjugated with a distinct oligonucleotide identification sequence on its surface that recognizes and hybridizes to a target gene (Table S2). We designed fluorescent Cy5-functionalized DNA secondary probes that hybridize to the target gene next to the identification sequence (Scheme 1, Table S2).

In our study, we designed panels of barcodes that are complementary to conserved, wildtype, and mutated sequences of SARS-CoV-2. The different barcode panels can detect SARS-CoV-2 variants because the secondary probe signal is different when binding to perfectly, partially, and noncomplementary sequences. Perfectly complementary sequences should yield the highest signal of the three. We quantify the secondary probe signal produced when a viral target binds to each of the different barcode sequences. We then compare secondary probe signals to create numerical ratios. We call this numerical value a “signal ratio”, and it allows us to genotype the virus. Further, we created a threshold that is normalized to 1.0, and values greater than 1.0 indicate a positive test and values less than 1.0 indicate a negative test. Table 2 provides the mathematical ratios for genotyping variants.

**Detecting Conserved SARS-CoV-2 Genes in Clinical Samples.** To genotype SARS-CoV-2 viral variants, we first check whether a sample was SARS-CoV-2-positive by examining our barcode results from the conserved gene regions and then examining our barcode results for mutations within the genes of interest (Schemes S1 and S2). We evaluated the feasibility of



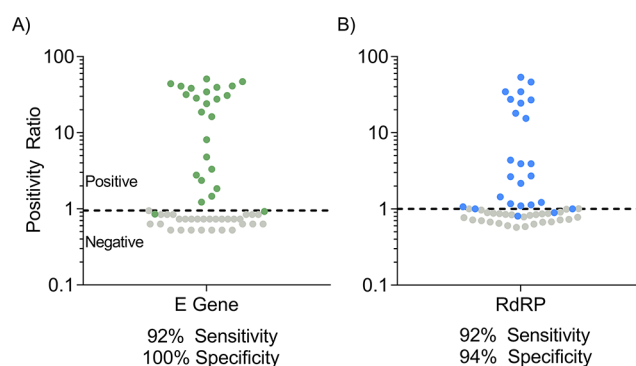
**Table 2. Description of Signal Ratio Calculation for Each Figure**

barcoding panel	signal ratio	target	barcodes used	figures
conserved gene panel	E gene signal ratio	E gene	E gene noncomplementary gene # 1	1A
conserved gene panel	RdRP gene signal ratio	RdRP gene	RdRP gene noncomplementary gene # 1	1B
H69/S70 deletion panel	E gene signal ratio	E gene	E gene noncomplementary gene # 2	S4
H69/S70 deletion panel	spike gene signal ratio	spike gene	H69 S70 spike wildtype H69 S70 spike deletion	S4
N501Y substitution panel	E gene signal ratio	E gene	E gene noncomplementary gene # 1	2B,E, 3A
N501Y substitution panel	spike gene signal ratio	spike gene	N501 spike wildtype N501Y spike mutant	2C,F, 3B

using signal ratios from a barcode panel to detect conserved SARS-CoV-2 genes in human clinical samples. It is important to evaluate the performance of our barcodes with human samples as the nonspecific binding of genomic DNA and varying viral loads of human samples affects the performance of diagnostics.<sup>26,27</sup> We designed a four-barcode panel to detect the highly conserved SARS-CoV-2 envelope (E) and RNA-dependent RNA polymerase (RdRP) genes (Table 1). Barcodes three and four were for two noncomplementary genes, which served as a reference signal and negative control, respectively. We detected the E and RdRP genes by creating normalized E or RdRP signal ratios using the first noncomplementary gene (Table 2). We tested 25 SARS-CoV-2-positive patient samples with RT-PCR cycle threshold (Ct) values ranging from 14 to 36 or viral loads ranging from  $7 \times 10^5$  copies per reaction to less than 1 copy per reaction.

We tested 30 SARS-CoV-2 negative clinical specimens with laboratory-confirmed influenza viruses A and B, coronaviruses (OC43, HKU1), and healthy human controls to evaluate the cross-reactivity and specificity of our barcodes. We extracted, reverse-transcribed, and isothermally amplified nucleic acids from each of the samples using the E and RdRP primers. We mixed the viral cDNA with the four-barcode SARS-CoV-2 panel, their respective secondary probes, and the noncomplementary gene targets. We then measured the barcode signal using flow cytometry. The normalized signal ratios ranged from 0.85 to 60 and from 0.77 to 59 for the E and RdRP genes, respectively (Figure 1A,B). We produced 92% sensitivity for both genes and 100 and 94% specificity for the E and RdRP gene, respectively, compared to RT-PCR (Figure 1, Table S3). Our results demonstrate that the E gene outperformed the RdRP gene; thus, we used the E gene to determine SARS-CoV-2 positivity in our other experiments. We demonstrated that our detection of conserved SARS-CoV-2 genes is not affected by the complexity of human samples and that our barcoding approach can be used to identify conserved genes.

**Genotyping a Nucleic Acid Deletion.** After evaluating the detection of conserved genes, we used our signal ratio approach to genotype SARS-CoV-2 samples that contained or did not contain deletion mutations (Scheme S3). We designed a barcode panel to simultaneously detect SARS-CoV-2 and genotype a six-nucleotide deletion (H69/S70) found in the Alpha (B.1.1.7) and Omicron (BA.1, BA.1.4, BA.1.5) spike genes (Table 1).<sup>28,29</sup> We confirmed our samples were positive for SARS-CoV-2 using the normalized E gene signal ratio (Table 2). Figure S4A shows that our signal ratio was consistently above

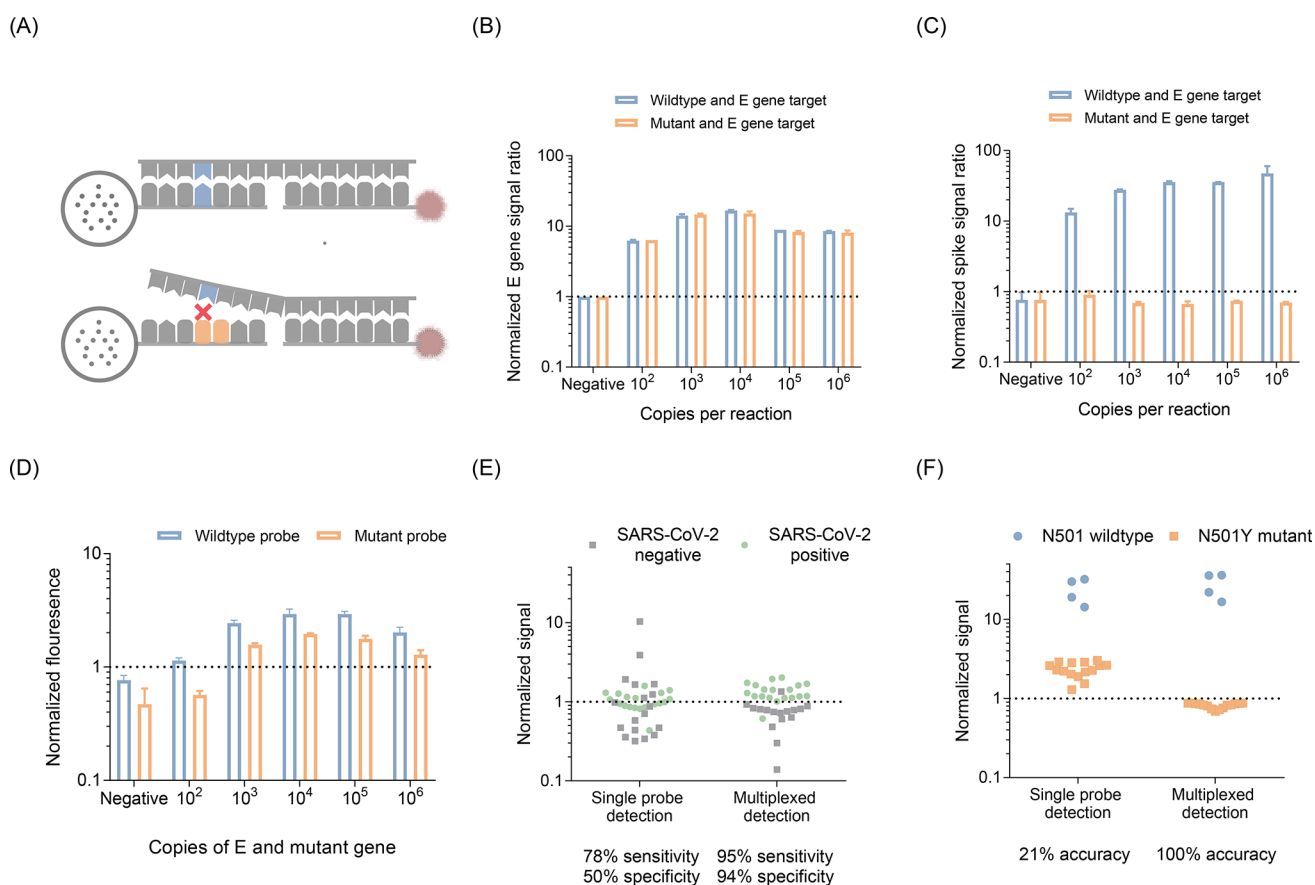


**Figure 1.** Detecting conserved SARS-CoV-2 genes in confirmed SARS-CoV-2 positive clinical samples. (A) E gene positivity ratio for the clinical samples detected using the conserved gene multiplexed barcode assay ( $n = 55$ ). The positivity ratio indicates the normalized E gene signal ratio, which is calculated by dividing the signal of the E gene barcode by the noncomplementary barcode. (B) RdRP gene positivity ratios for the clinical samples detected using the conserved gene multiplexed barcode assay. The positivity ratio indicates the normalized RdRP gene signal ratio, which is calculated by dividing the signal of the RdRP gene barcode by the noncomplementary barcode. Sensitivity and specificity values calculated by comparing results to qRT-PCR.

1.0 for the E gene when we used our barcode panel to detect SARS-CoV-2 samples with or without the six-nucleotide spike gene deletion. We demonstrate that the detection of the conserved SARS-CoV-2 E gene is not cross-reactive with either the spike deletion or wildtype gene. We can identify the conserved E gene in our multiplex barcode panel.

Once we evaluated the detection of the conserved E gene in our panel, we used our barcode panel to detect the H69/S70 spike gene deletion by calculating a “spike gene signal ratio” (Scheme S3, Tables 1 and 2). The wildtype and deletion spike gene barcodes were designed to be perfectly complementary to their respective spike sequences. The spike gene signal ratio reflects the preferential binding of a sequence to the wildtype or deletion barcode, thereby indicating the genotype. The normalized spike gene ratios were below 1.0 for E gene-positive samples that contained the deletion spike gene (Figure S4). The wildtype spike barcode produced less signal than the deletion spike barcode in the presence of a spike deletion target. Each deletion spike gene target was correctly genotyped as having the H69/S70 deletion. In contrast, the normalized spike gene ratios for the E gene positive samples with the wildtype spike gene were above 1.0. Each wildtype spike gene target was correctly genotyped as missing the H69/S70 deletion. The difference in hybridization between the deletion or wildtype barcode identification sequences and the spike target affected the resulting signal ratio. We simultaneously detected SARS-CoV-2 and the presence of a genetic deletion by comparing the signal produced in our four-barcode panel.

**Genotyping a Single Nucleotide Substitution Using Multiple Barcodes.** After using our barcoding approach to detect a genetic deletion, we designed a barcode panel to genotype the N501Y single nucleotide mutation in the spike gene of circulating SARS-CoV-2 variants (Scheme S4, Figure 2, Table 1).<sup>28,29</sup> Our barcode panel simultaneously detected SARS-CoV-2 and the N501Y spike gene mutation in a single test. We confirmed that the samples were SARS-CoV-2-positive using the normalized E gene signal ratio. To test the limit of E gene detection in our N501Y barcode panel, we conducted two sets of experiments detecting amplified E gene in the presence of



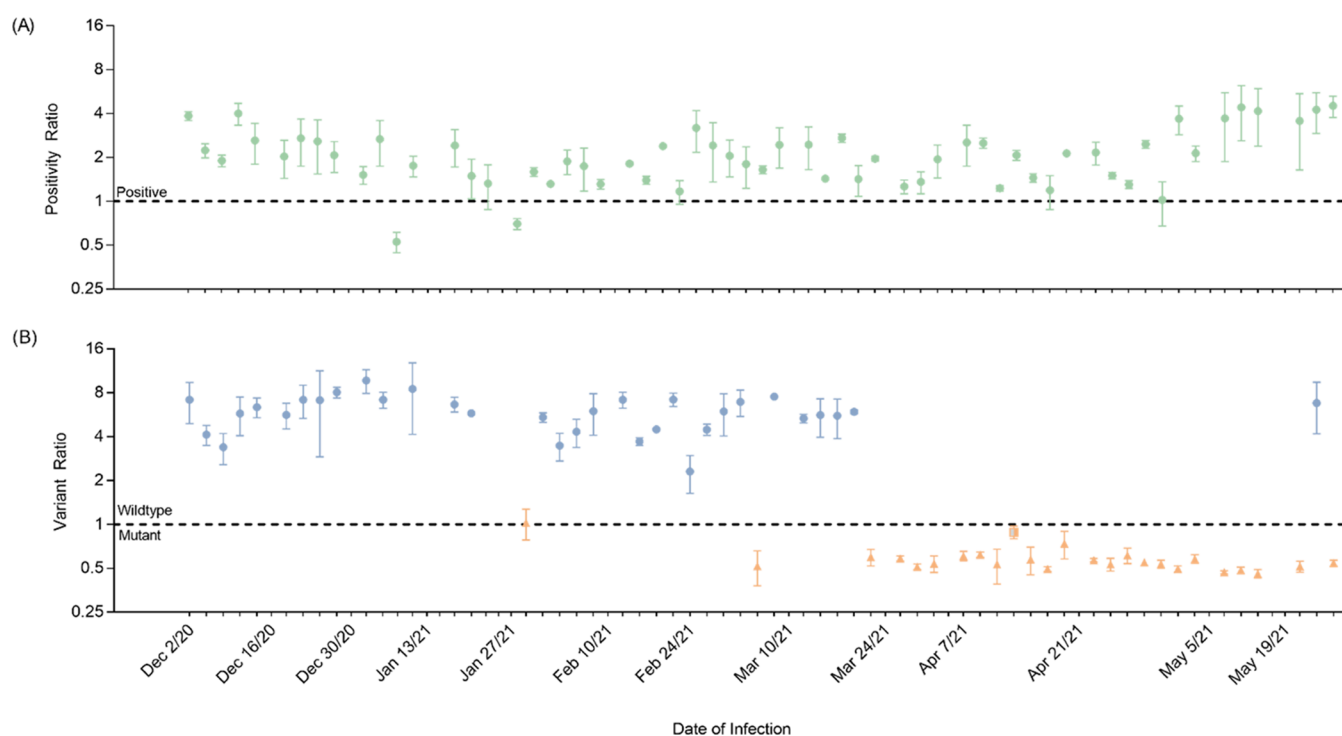
**Figure 2.** Detecting the N501Y single nucleotide substitution in the SARS-CoV-2 spike gene. (A) This scheme demonstrates that a single nucleotide mutation will bind strongly to a complementary probe and weakly to a partially complementary probe. (B) Normalized E gene ratios are used to detect the presence of the E gene when mixed with either wildtype or mutated spike gene. Any E gene ratio value above 1.0 is classified as SARS-CoV-2 positive. (C) Normalized spike gene ratios used to genotype the N501Y mutation from samples containing either spike wildtype and E gene or spike mutant and E gene. Any spike signal ratio above 1.0 is classified as a wildtype sample. (D) Normalized fluorescence signal generated from a single-probe genotyping system using the spike wildtype barcode or spike mutant barcode in the presence of mutant spike and E gene. Any value above the background (dashed line) is classified as a wildtype sample. (E) Normalized signal-to-cutoff values for SARS-CoV-2 E gene detection from SARS-CoV-2 qRT-PCR positive human samples. The E gene was detected using the E gene barcode alone or a multiplexed barcode approach. Any value above the threshold (dashed line) is considered SARS-CoV-2-positive with its respective assay. (F) Normalized signal-to-cutoff values for genotyping the SARS-CoV-2 N501Y spike mutation from clinical samples. The N501Y mutation was genotyped using only the wildtype barcode or our multiplex barcode approach. Any value above the threshold (dashed line) is genotyped as N501Y wildtype using the respective assay.

amplified wildtype spike gene or amplified N501Y mutant spike gene from 10<sup>6</sup> to 10<sup>2</sup> copies per reaction. We detected down to 100 copies of E gene in the presence of either wildtype or mutant spike DNA (Figure 2B). We demonstrated that the detection of amplified E gene was not cross-reactive with either amplified wildtype or mutant spike genes in our barcode assay. We can identify E gene in the presence of single nucleotide N501Y variant sequences. Once we evaluated the detection of the conserved E gene in our panel, we genotyped the N501Y substitution for SARS-CoV-2 by calculating a “spike gene signal ratio” using our barcode panel (Tables 1 and 2). We genotyped the N501Y substitution for SARS-CoV-2 positive samples. The normalized spike gene signal ratios were below 1.0 for 10<sup>6</sup>–10<sup>2</sup> copies of mixed E and mutant spike gene targets and above 1.0 for 10<sup>6</sup>–10<sup>2</sup> copies of mixed E and wildtype spike gene targets (Figure 2C). We demonstrated that we could use our barcode approach to identify and genotype synthetic samples with the N501Y single nucleotide substitution.

Next, we showed that our barcode and signal ratio approach improved the genotyping of viral sequences compared to single-probe genotyping. For single-probe genotyping, we incubated

one sample with one spike barcode rather than with multiple barcodes. Figure 2D shows the mutant spike gene incorrectly appeared positive using both the wildtype and mutant barcodes. The mutant gene produced a positive signal in the presence of either barcode, demonstrating the inability to genotype with a single-probe system. This cross-reactivity is a problem for single barcode detection systems, but we leveraged this phenomenon to produce signal ratios that identify the N501Y single nucleotide mutation in our samples. Unlike single-probe systems, our multiplex barcoding and signal ratio approach correctly genotyped samples with the N501Y single nucleotide substitution.

**Clinically Validating the SARS-CoV-2 Single Nucleotide Substitution Genotyping Panel.** Once we had evaluated that our barcoding approach could genotype the N501Y spike gene variant in synthetic samples, we used our barcode panel to genotype the N501Y in SARS-CoV-2-positive clinical specimens. We collected 19 patient samples containing wildtype (N501),  $\alpha$ ,  $\beta$ , or  $\gamma$  SARS-CoV-2 variants, then reverse-transcribed, and isothermally amplified the spike and E gene regions. The barcode panel achieved 95% sensitivity in detecting



**Figure 3.** Emergence of SARS-CoV-2 variants in Ontario, Canada. SARS-CoV-2 samples were tested by qRT-PCR and using the N501Y signal ratio panel. Panel (A) indicates SARS-CoV-2 positivity, and panel (B) indicates whether samples are spike wildtype (N501) or spike mutant (N501Y) using the multiplex N501Y signal ratio panel. Variant ratios were calculated for samples positive for SARS-CoV-2 (green circles). In late 2020, most samples were wildtype (blue circles). In the spring of 2021, most samples were  $\alpha$  (orange triangles), and a fraction of samples were  $\gamma$  (orange-gray boxes). The position of the symbol relative to the threshold indicates signal ratio results, and the color of the symbol indicates qRT-PCR results.

SARS-CoV-2 and 100% accuracy in genotyping the N501Y wildtype or mutant spike gene (Figure 2E,F). We correctly identified the presence or absence of spike gene mutations in all clinical samples. Our approach showed 94% specificity and low cross-reactivity to adenovirus, other human coronaviruses, enterovirus, influenza viruses, norovirus, parainfluenza viruses, respiratory syncytial virus, and rhinoviruses (Figure 2E, Table S4).

We compared the performance to a single-probe genotyping format. We demonstrated that a single barcode performed worse than our multiplex signal ratio approach. We used only a single barcode to detect the samples and the single barcode approach produced 78% sensitivity, 50% specificity, and 21% accuracy (Figure 2E,F). The single nucleotide N501Y substitution has a small effect on the binding between the wildtype barcode and an N501Y mutant sequence. This effect can be leveraged using a multiplexed detection scheme but is not tolerated in a single-probe genotyping format.

**Tracking the Evolution of the N501Y SARS-CoV-2 Spike Gene Mutation.** Once we had established that our barcoding approach was highly accurate, we applied our N501Y barcode panel to track the transmission of N501Y SARS-CoV-2 variants in Ontario, Canada. SARS-CoV-2 variants containing the N501Y mutation are more transmissible than wildtype SARS-CoV-2.<sup>28,30</sup> We evaluated samples from patients infected between December 2020 and May 2021 (Figure 3A). Green circles show the SARS-CoV-2-positive samples as determined by RT-PCR. Samples with signal ratios above the dotted line (positivity threshold) were considered positive for SARS-CoV-2 using the N501Y signal ratio system. We identified 97% of samples as SARS-CoV-2-positive using the N501Y probe panel

(Table S4, Table S5) using the E gene to noncomplementary gene #1 signal ratio (Table 2).

Once we had confirmed that our samples were SARS-CoV-2-positive, we interpreted the genotyping results using the N501 spike gene barcode ratios. Blue circles represent wildtype samples without the N501Y mutation (Figure 3B), orange triangles represent the SARS-CoV-2  $\alpha$  variant, and the orange box represents the  $\beta$  variant as confirmed using RT-PCR. Samples with a signal ratio above and below the dotted line were considered wildtype and mutant spike, respectively. We observed that from December 2020 to February 2021, only a few SARS-CoV-2 samples contained the more transmissible N501Y mutation. This was followed by a sharp increase in N501Y mutants in April 2021. Only 10% of samples in April were wildtype, and 90% contained the N501Y mutation. The observed increase in N501Y mutations aligned with the rise in SARS-CoV-2 cases and N501Y mutants recorded by provincial public health laboratories.<sup>31</sup> Our barcoding and signal ratio approach can track the evolution of viral variants in a population.

## CONCLUSIONS

Quantum dot barcoding technology has been developing for the last two decades. While the basic concepts were proposed in the early 2000s, it takes significant efforts to develop such technology for commercial and broad use. The need for translation led researchers to focus their efforts in optimizing barcode synthesis and surface functionality,<sup>32–34</sup> and evaluating performance against different targets.<sup>25,35,36</sup> These studies established the components for using quantum dot barcodes. The novelty of this study is the advancement of the barcode technology for genotyping infectious diseases by combining multiplexing with ratiometric analysis. It overcomes the



limitations of genotyping technologies for the rapid monitoring of infectious diseases. Our approach produced a 5 times improvement in the detection accuracy of detecting SARS-CoV-2 N501 variants compared to the conventional single-probe detection scheme. While we used quantum dots for the barcoding technology, our approach can be adapted for other barcoding and multiplex schemes (Luminex arrays, CRISPR-based diagnostics, and multi-probe qPCR). Broadly, simple genotyping tools for diseases are critical to preventing the spread of diseases, and our approach is a step toward the engineering of simpler detection systems for monitoring the spread of pathogen variants.

Compared to the current state-of-the-art and conventional genotyping systems, our ratiometric barcoding concept has advantages: (1) Lenient probe specificity criteria where probes that do not exclusively bind to one sequence can be used for detection, which is unlike single-probe detection systems that require highly specific probes.<sup>17,19</sup> (2) Avoids the need for complex analytical or cross-reactivity evaluation for each probe. Each probe does not need to be tested for cross-reactivity with other targets, as the ratios identify targets even in the presence of cross-reactivity. (3) Our approach is adaptable to technologies with high cross-reactivity, such as protein detection (e.g., Zika and Dengue viruses).<sup>37–39</sup> However, our barcoding system cannot determine unknown variants, but it will indicate that a virus is mutating. This result will lead researchers to start to do sequencing to identify the mutations. Current genome sequencing is impractical for monitoring unknown variants. Our barcoding approach will require less infrastructure and faster turnaround time when coupled with an automated device capable of extraction and detection of viral targets. Our lab is currently working on such a device. This device would be portable and relay information to centralized facilities.

Our ratiometric analysis can be scaled to design panels with hundreds of probes to create massively multiplexed point-of-care diagnostic technologies to monitor mutations in a population. We can design quantum dot barcodes to detect over 100 probes, which can be combined to generate hundreds of signal ratios.<sup>22</sup> Each signal ratio is a data point. Adding a single target sequence to the probe panel generates hundreds of data points that can be combined to describe the sequence identity. These detailed sequence descriptions can track changes in the genome sequence of a population. For example, numerous probes can be designed to detect a rapidly mutating gene linked to virus transmissibility. Probes will have similar sequences and bind to similar regions. The data from individual probes would be inconsequential, but the compilation of signal ratio data would indicate sequence identity and help predict transmissibility. Our group has demonstrated point-of-care detection using smartphone-based quantum dot barcode assays.<sup>35,36</sup> These 100 probe panels could be integrated with smartphone point-of-care detection and allow public health agencies to rapidly monitor the evolution of new or existing infectious diseases.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.2c04630>.

Workflow for determining E gene positivity, RDRP positivity H60/S70 genotype, and N501Y genotype (Schemes S1–S4); quantum dot images and flow analysis

(Figures S1 and S2); barcode limit of detection (Figure S3); quantum dot barcode synthesis volumes (Table S1); oligonucleotides used in the study (Table S2); and clinical sample information (Tables S3, S4) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Warren C. W. Chan** — Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada; Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada; Department of Chemical Engineering, University of Toronto, Toronto, Ontario M5S 3E5, Canada; [orcid.org/0000-0001-5435-4785](https://orcid.org/0000-0001-5435-4785); Email: [warren.chan@utoronto.ca](mailto:warren.chan@utoronto.ca)

### Authors

**Hannah N. Kozlowski** — Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada; Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada; MD/PhD Program, Temerty Faculty of Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada

**Ayden Malekjahani** — Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada; Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

**Vanessa Y. C. Li** — Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada; Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

**Ayokunle A. Lekuti** — Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada; Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

**Stephen Perusini** — Public Health Ontario, Toronto, Ontario M5G 1V2, Canada

**Natalie G. Bell** — Department of Laboratory Medicine and Pathobiology and Faculty of Medicine, University of Toronto, Toronto, Ontario M5S 1A1, Canada; Biological Sciences, Sunnybrook Research Institute, Toronto, Ontario M4N 3M5, Canada

**Veronique Voisin** — Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

**Delaram Pouyababar** — Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

**Shraddha Pai** — Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

**Gary D. Bader** — Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada; Department of Molecular Genetics and Department of Computer Science, University of Toronto, Toronto, Ontario M5S 3E1, Canada; Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario M5G 2C1, Canada; The Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario M5G 1X5, Canada

**Samira Mubareka** – Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario M5S 1A1, Canada; Biological Sciences, Sunnybrook Research Institute, Toronto, Ontario M4N 3M5, Canada

**Jonathan B. Gubbay** – Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario M5S 1A1, Canada; Public Health Ontario, Toronto, Ontario M5G 1V2, Canada

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.analchem.2c04630>

## Author Contributions

<sup>‡</sup>H.N.K. and A.M. contributed equally to this work.

## Notes

The authors declare the following competing financial interest(s): Warren C.W. Chan declares that he co-founded Luna Nanotech and consult for Foresight Capital and Cystic Fibrosis Foundation.

## ACKNOWLEDGMENTS

The authors acknowledge the Canadian Institute of Health Research (CIHR, FDN-159932; MOP-130143), Natural Sciences and Engineering Research Council of Canada (NSERC, 2015-06397), Canadian Research Chairs program (950-223824), Collaborative Health Research Program (CPG146468), and Canadian Cancer Society (705185-1) for funding support. We thank Ruth Isserlin, Shirley Hui, Maria Abou Chakra and Zoe Clarke for help developing the original bioinformatics pipeline. H.N.K. thanks Paul Cadario, Ruggles, the Vanier Canadian Graduate Scholarship from CIHR, the McLaughlin Institute, and the University of Toronto MD/PhD program. A.M. thanks the Postgraduate Scholarship from NSERC. V.Y.C.L. thanks the Barbara and Frank Milligan Graduate Fellowship, Doctoral Canada Graduate Scholarship from NSERC, and Nanomedicine Innovation Network Graduate award. A.A.L. thanks the Sally and Paul Wang Distinguished Graduate Scholarship and the Barbara and Frank Milligan Graduate Fellowship.

## REFERENCES

- (1) Hirschhorn, J. N.; Lohmueller, K.; Byrne, E.; Hirschhorn, K. *Genet. Med.* **2002**, *4*, 45–61.
- (2) Zhou, S.; Butler-Laporte, G.; Nakanishi, T.; Morrison, D. R.; Afilalo, J.; Afilalo, M.; Laurent, L.; Pietzner, M.; Kerrison, N.; Zhao, K.; Brunet-Ratnasingham, E.; Henry, D.; Kimchi, N.; Afrasiabi, Z.; Rezk, N.; Bouab, M.; Petitjean, L.; Guzman, C.; Xue, X.; Tselios, C.; Vulesevic, B.; Adeleye, O.; Abdullah, T.; Almamlouk, N.; Chen, Y.; Chassé, M.; Durand, M.; Paterson, C.; Normark, J.; Frithiof, R.; Lipcsey, M.; Hultström, M.; Greenwood, C. M. T.; Zeberg, H.; Langenberg, C.; Thysell, E.; Pollak, M.; Mooser, V.; Forgetta, V.; Kaufmann, D. E.; Richards, J. B. *Nat. Med.* **2021**, *27*, 659–667.
- (3) Kramvis, A.; Kew, M.; François, G.; Hepatitis, B. *Virus Genotype Vaccine* 2005.
- (4) Eroglu, C.; Leblebicioglu, H.; Gunaydin, M.; Turan, D.; Sunbul, M.; Esen, S.; Sanic, A. *J. Virol. Methods* **2004**, *119*, 183–187.
- (5) Zhang, D.; Ma, S.; Zhang, X.; Zhao, H.; Ding, H.; Zeng, C. *BMC Infect. Dis.* **2010**, *10*, No. 271.
- (6) Call, D. R.; Brockman, F. J.; Chandler, D. P. *Int. J. Food Microbiol.* **2001**, *67*, 71–80.
- (7) Kwok, P. Y. *Annu. Rev. Genomics Hum. Genet.* **2001**, *2*, 235–258.
- (8) Armour, J. A. L.; Barton, D. E.; Cockburn, D. J.; Taylor, G. R. *Hum. Mutat.* **2002**, *20*, 325–337.
- (9) Behjati, S.; Tarpey, P. S. *Arch. Dis. Child. Educ. Pract. Ed.* **2013**, *98*, 236–238.
- (10) Alekseyev, Y. O.; Fazeli, R.; Yang, S.; Basran, R.; Maher, T.; Miller, N. S.; Remick, D. *Acad. Pathol.* **2018**, *5*, No. 2374289518766521.
- (11) Goodwin, S.; McPherson, J. D.; McCombie, W. R. *Nat. Rev. Genet.* **2016**, *17*, 333–351.
- (12) Khodakov, D.; Li, J.; Zhang, J. X.; Zhang, D. Y. *Nat. Biomed. Eng.* **2021**, *5*, 702–712.
- (13) Hert, D. G.; Fredlake, C. P.; Barron, A. E. *Electrophoresis* **2008**, *29*, 4618–4626.
- (14) Why COVID-19 Testing is the Key to Getting Back to Normal; National Institute on Aging, <https://www.nia.nih.gov/news/why-covid-19-testing-key-getting-back-normal> (accessed Dec 12, 2022).
- (15) Peeling, R. W.; Heymann, D. L.; Teo, Y.-Y.; Garcia, P. J. *Lancet* **2022**, *399*, 757–768.
- (16) Malekjhani, A.; Sindhiani, S.; Syed, A. M.; Chan, W. C. W. *Acc. Chem. Res.* **2019**, *52*, 2406–2414.
- (17) Mouritzen, P.; Nielsen, A. T.; Pfundheller, H. M.; Choleva, Y.; Kongsbak, L.; Møller, S. *Expert Rev. Mol. Diagn.* **2003**, *3*, 27–38.
- (18) Koshkin, A. A.; Singh, S. K.; Nielsen, P. *Tetrahedron* **1998**, *54*, 3607–3630.
- (19) Johnson, M. P.; Haupt, L. M.; Griffiths, L. R. *Nucleic Acids Res.* **2004**, *32*, No. e55.
- (20) Hashim, H. O.; Al-Shuhaib, M. B. *J. Appl. Biotechnol. Rep.* **2019**, *6*, 137–144.
- (21) Gaudet, M.; Fara, A.-G.; Beritognolo, I.; Sabatti, M. *Methods Mol. Biol.* **2009**, *578*, 415–424.
- (22) Fournier-Bidoz, S.; Jennings, T. L.; Klostranec, J. M.; Fung, W.; Rhee, A.; Li, D.; Chan, W. C. W. *Angew. Chem., Int. Ed.* **2008**, *47*, 5577–5581.
- (23) Corman, V. M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D. K.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M. L.; Mulders, D. G.; Haagmans, B. L.; van der Veer, B.; van den Brink, S.; Wijsman, L.; Goderski, G.; Romette, J.-L.; Ellis, J.; Zambon, M.; Peiris, M.; Goossens, H.; Reusken, C.; Koopmans, M. P.; Drosten, C. *Euro Surveill.* **2020**, *25*, No. 2000045.
- (24) Li, D.; Zhang, J.; Li, J. *Theranostics* **2020**, *10*, 7150–7162.
- (25) Kim, J.; Biondi, M. J.; Feld, J. J.; Chan, W. C. W. *ACS Nano* **2016**, *10*, 4742–4753.
- (26) Zhou, W.; Yang, S.; Wang, P. G. *Bioanalysis* **2017**, *9*, 1839–1844.
- (27) Chiu, M. L.; Lawi, W.; Snyder, S. T.; Wong, P. K.; Liao, J. C.; Gau, V. *J. Lab. Autom.* **2010**, *15*, 233–242.
- (28) Harvey, W. T.; Carabelli, A. M.; Jackson, B.; Gupta, R. K.; Thomson, E. C.; Harrison, E. M.; Ludden, C.; Reeve, R.; Rambaut, A.; Peacock, S. J.; Robertson, D. L. *Nat. Rev. Microbiol.* **2021**, *19*, 409–424.
- (29) Lippi, G.; Mattiuzzi, C.; Henry, B. M. *Diagnosis* **2022**, *9*, 11–17.
- (30) Escalera, A.; Gonzalez-Reiche, A. S.; Aslam, S.; Mena, I.; Laporte, M.; Pearl, R. L.; Fossati, A.; Rathnasinghe, R.; Al-Shammari, H.; Van De Guchte, A.; Farrugia, K.; Qin, Y.; Bou-Haddou, M.; Kehrer, T.; Zuliani-Alvarez, L.; Meekins, D. A.; Balaraman, V.; McDowell, C.; Richt, J. A.; Bajic, G.; Sordillo, E. M.; Dejoze, M.; Zwaka, T. P.; Krogan, N. J.; Simon, V.; Albrecht, R. A.; Van Bakel, H.; Garcia-Sastre, A.; Aydiello, T. *Cell Host Mi-crobe* **2022**, *30*, 373–387.
- (31) Ontario COVID-19 Data Tool; Public Health Ontario. <https://www.publichealthontario.ca/en/data-and-analysis/infectious-disease/covid-19-data-surveillance/covid-19-data-tool?tab=trends> (accessed Dec 12, 2022).
- (32) Rauf, S.; Glidle, A.; Cooper, J. M. *Adv. Mater.* **2009**, *21*, 4020–4024.
- (33) Klostranec, J. M.; Xiang, Q.; Farcas, G. A.; Lee, J. A.; Rhee, A.; Lafferty, E. I.; Perrault, S. D.; Kain, K. C.; Chan, W. C. W. *Nano Lett.* **2007**, *7*, 2812–2818.
- (34) Giri, S.; Sykes, E. A.; Jennings, T. L.; Chan, W. C. W. *ACS Nano* **2011**, *5*, 1580–1587.
- (35) Ming, K.; Kim, J.; Biondi, M. J.; Syed, A.; Chen, K.; Lam, A.; Ostrowski, M.; Rebbapragada, A.; Feld, J. J.; Chan, W. C. W. *ACS Nano* **2015**, *9*, 3060–3074.
- (36) Zhang, Y.; Malekjhani, A.; Udugama, B. N.; Kadhiresan, P.; Chen, H.; Osborne, M.; Franz, M.; Kucera, M.; Plenderleith, S.; Yip, L.;



Bader, G. D.; Tran, V.; Gubbay, J. B.; McGeer, A.; Mubareka, S.; Chan, W. C. W. *Nano Lett.* **2021**, *21*, S209–S216.

(37) Kam, Y.-W.; Lee, C. Y.; Teo, P.; Howland, T.-H.; Am-Run, S. W.; Lum, S. N.; See, F.-M.; Kng, P.-R. N. Q.; Huber, R. G.; Xu, M.-H.; Tan, H.-L.; Choo, A.; Maurer-Stroh, S.; Ginhoux, F.; Fink, K.; Wang, C.-I.; Ng, L. F.; Rénia, L. *JCI Insight* **2017**, *2*, No. e92428.

(38) Felix, A. C.; Souza, N. C. S.; Figueiredo, W. M.; Costa, A. A.; Inenami, M.; da Silva, R. M. G.; Levi, J. E.; Pannuti, C. S.; Romano, C. M. *J. Med. Virol.* **2017**, *89*, 1477–1479.

(39) Priyamvada, L.; Quicke, K. M.; Hudson, W. H.; Onlamoon, N.; Sewatanon, J.; Edupuganti, S.; Pattanapanyasat, K.; Chokephaibulkit, K.; Mulligan, M. J.; Wilson, P. C.; Ahmed, R.; Suthar, M. S.; Wrammert, J. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 7852–7857.

(40) Rejali, N. A.; Moric, E.; Wittwer, C. T. *Clin. Chem.* **2018**, *64*, 801–809.

(41) Naiser, T.; Ehler, O.; Kayser, J.; Mai, T.; Michel, W.; Ott, A. *BMC Biotechnol.* **2008**, *8*, 48.

## Recommended by ACS

### Enhancing Efficiency and Stability of Perovskite Solar Cells via Photosensitive Molecule-Assisted Defect Passivation

Mingguang Li, Runfeng Chen, *et al.*

MARCH 31, 2023  
ACS APPLIED ENERGY MATERIALS

READ 

### Experimental and DFT Studies on Electrochemical Performances of Ester-Containing Vinylimidazolium Ionic Liquids: Effect of the Ester Substituent and Anion

Yupei Deng, Linghua Zhuang, *et al.*

MARCH 30, 2023  
JOURNAL OF CHEMICAL & ENGINEERING DATA

READ 

### Computation-Assisted Design of ssDNA Framework Nanorobots for Cancer Logical Recognition, Toehold Disintegration, Visual Dual-Diagnosis, and Synergistic Th...

Xiaotong Shen, Na Na, *et al.*

MARCH 31, 2023  
ANALYTICAL CHEMISTRY

READ 

### Thermodynamic Study of the Activity of Cadmium Chloride in Aqueous Solutions at Various Temperatures

Brahim Makka, Asmaa Benbiyi, *et al.*

MARCH 30, 2023  
JOURNAL OF CHEMICAL & ENGINEERING DATA

READ 

Get More Suggestions >