Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Real-time allelic assays of SARS-CoV-2 variants to enhance sewage surveillance

Xiaoqing Xu a, Yu Deng a, Jiahui Ding a, Xiawan Zheng a, Shuxian Li a, Lei Liu a, Ho-kwong Chui b, Leo L.M. Poon c, Tong Zhang a,*

a Environmental Microbiome Engineering and Biotechnology Lab, Center for Environmental Engineering Research, Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China
b Environmental Protection Department, The Government of Hong Kong SAR, Tamar, Hong Kong SAR, China
c School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Sassoon Road, Hong Kong SAR, China

ARTICLE INFO
Keywords:
SARS-CoV-2 Variants Allele-specific RT-qPCR Real-time Sewage surveillance

ABSTRACT
To effectively control the ongoing outbreaks of fast-spreading SARS-CoV-2 variants, there is an urgent need to add rapid variant detection and discrimination methods to the existing sewage surveillance systems established worldwide. We designed eight assays based on allele-specific RT-qPCR for real-time allelic discrimination of eight SARS-CoV-2 variants (Alpha, Beta, Gamma, Delta, Omicron, Lambda, Mu, and Kappa) in sewage. In silico analysis of the designed assays for identifying SARS-CoV-2 variants using more than four million SARS-CoV-2 variant sequences yielded ~100% specificity and >90% sensitivity. All assays could sensitively discriminate and quantify target variants at levels as low as 10 viral RNA copy/µL with minimal cross-reactivity to the corresponding nontarget genotypes, even for sewage samples containing mixtures of SARS-CoV-2 variants with differential abundances.

Integration of this method into the routine sewage surveillance in Hong Kong successfully identified the Beta variant in a community sewage. Complete concordance was observed between the results of viral whole-genome sequencing and those of our novel assays in sewage samples that contained exclusively the Delta variant discharged by a clinically diagnosed COVID-19 patient living in a quarantine hotel. Our assays in this method also provided real-time discrimination of the newly emerging Omicron variant in sewage two days prior to clinical test results in another quarantine hotel in Hong Kong. These novel allelic discrimination assays offer a rapid, sensitive, and specific way for detecting multiple SARS-CoV-2 variants in sewage and can be directly integrated into the existing sewage surveillance systems.

1. Introduction
The continuous emergence and circulation of different SARS-CoV-2 variants carrying multiple mutations continue to pose new challenges to public health (Challen et al., 2021). As of November 2021, the World Health Organization (WHO) listed five SARS-CoV-2 variants of concern (VOCs), namely, Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529), and two variants of interest (VOIs), Lambda (C.37) and Mu (B.1.621). Notably, the Delta variant, which was first recorded in late 2020, has spread across 104 countries, and outcompeted the lineages of Alpha and Kappa (B.1.617.1, formerly a VOI, currently a variant under monitoring (VUM)). Rapid surveillance of SARS-CoV-2 variants in sewage as an early warning tool is urgently needed, especially when global populations are being vaccinated against specific SARS-CoV-2 strains.

Sewage surveillance offers a cost-effective strategy for monitoring disease prevalence at the population level. For COVID-19, sewage surveillance has been demonstrated to be useful in providing early warning signals for virus emergence and reintroduction (Ahmed et al., 2020; Xu et al., 2021), tracking epidemic trends (D'Aoust et al., 2021; Peccia et al., 2020), and identifying hidden cases in communities (Deng et al., 2022a; Mallaparty, 2020; Medema et al., 2020) and in transmission hotspots such as student dormitories (Gibas et al., 2021; Korfmancher et al., 2021; Zhang, 2022). Many countries and regions worldwide have established systematic sewage surveillance for COVID-19, including the United States (Keshaviah et al., 2021), the Netherlands (Kitajima et al.,...
2020), Australia (Ahmed et al., 2020), and Hong Kong (Deng et al., 2022b; Xu et al., 2021), etc. Since the end of 2020, several mutations emerging in the SARS-CoV-2 genome have been reported to be associated with increased virulence, infectivity and transmission rate (Chen et al., 2020; Korber et al., 2020; Team, 2021) and reduced sensitivity to antibodies (Gaebele et al., 2021; Wang et al., 2021). In principle, sewage testing can be adapted to identify SARS-CoV-2 variants that carry these concerning mutations.

However, the detection methods currently used in sewage surveillance are mainly based on reverse transcription-quantitative polymerase chain reaction (RT-qPCR) using primer-probe sets targeting convered regions of the viral genome (Xu et al., 2022), such as the N1, N2, and E genes, and therefore cannot discriminate SARS-CoV-2 variant types. Early efforts were made to identify SARS-CoV-2 variants in sewage using a sequencing-based approach (Agrawal et al., 2022; Crits-Christoph et al., 2021; Izquierdo-Lara et al., 2021; Martin et al., 2020). However, it is challenging to apply this approach to sewage samples due to the complexity of the sewage matrix, the relatively low viral titers, fragmented viral genomes, and mixtures of viral variants in these samples. From the perspective of informing policy-making agencies, one of the unique strengths of sewage surveillance is that it can offer a lead time of several days for the detection of infection(s) in a community compared with testing of individuals (Wu et al., 2022). However, the sequencing-based approach is a time-consuming and expensive method and requires considerable optimization for sewage samples to cope with the rapid turnaround and extensive surveillance needed for policymakers.

Several PCR-based approaches have been developed for the rapid discrimination of SARS-CoV-2 variants in clinical samples, such as the RT-PCR melt temperature assay (Banada et al., 2021) and multiplex qPCR target failure screening (Vogels et al., 2021). The former method requires a clear difference in the annealing temperatures between the target mutation sequence and its opposite genotype, which could be affected by any additional mutations located in the target sequence (Gazali et al., 2021). The multiplex qPCR target failure screening is based on deletions in the genome of SARS-CoV-2, and does not have good sensitivity and specificity since the rapid emergence of new mutations is the major concern. Additionally, it cannot be applied to detect mutations caused by substitutions (Vogels et al., 2021). Adapting PCR-based approaches for variant detection in sewage requires specificity and sensitivity validation at the same concentration level as in sewage samples. Several studies using probe-based RT-qPCR to detect SARS-CoV-2 variants (Peterson et al., 2022; Yaniv et al., 2021), and commercial kits (e.g., TaqPath Thermo Fisher), have recently emerged. However, a major disadvantage of the probe-based approach is increased design complexity and cost (Kalender et al., 2021; Lefever et al., 2019), and for commercial kits, assay details are not available. As an alternative, the cost-effective allele-specific (AS) RT-qPCR-based method, has shown the potential to reliably identify the Alpha variant (Lee et al., 2021). This method was developed for molecular genotyping, which allows the direct detection of alleles of the wild type and its mutant in the genomic sequence with the same general RT-qPCR workflow by using two different primers complementary to the wild-type and mutant alleles specifically and one common primer and a common probe for both alleles (Wu et al., 1989). To ensure high levels of specificity, a mismatch can be introduced near the 3’ end of allele-specific primers, which makes the nontarget sequences more refractory to primer extension (O’Meara et al., 2002). Recent studies reported the use of the AS RT-qPCR approach to detect the Alpha variant in sewage samples, based on either three characteristic mutation sites in the S gene (Lee et al., 2021) or one signature mutation site in the N gene (Graber et al., 2021). However, currently, there is no comprehensive method using AS RT-qPCR assays for the determination of multiple SARS-CoV-2 variants in sewage samples; more specifically, there is no method yet available for discriminating variants with increased transmissibility and newly emerging variants.

In this study, a set of novel assays based on AS RT-qPCR were developed to identify different SARS-CoV-2 variants in sewage samples. The eight assays were composed of different combinations of 12 spike protein mutation sites to distinguish eight SARS-CoV-2 variants, including five VOCs, two VOIs, and one VUM as designated by the WHO (Konings et al., 2021; Parums, 2021). In silico specificity evaluation for the target mutation sites of the designed primer-probe sets was conducted with more than four million SARS-CoV-2 genomic sequences in the Global Initiative on Sharing All Influenza Data (GISAID) database. The analytic performance of the assays and the sewage matrix effects were assessed with the RNA and DNA controls. The quantitative accuracy of these assays was further examined in mock sewage samples containing a mixture of SARS-CoV-2 variants. Incorporation of these assays into the routine COVID-19 sewage testing in Hong Kong strengthened the surveillance by providing real-time variant genotyping of the detected SARS-CoV-2, as exemplified by three use cases in identifying the Beta variant in community sewage and documenting Delta and Omicron importation at quarantine hotels. Further viral whole-genome sequencing using sewage samples confirmed the identification accuracy of our assays. These assays could be directly implemented in the existing sewage surveillance systems to detect circulating variants in sewage.

2. Materials and methods

2.1. Primer and probe design principles

The primers and probes targeting 12 spike protein mutation sites to distinguish eight SARS-CoV-2 variants (Fig. S1) were designed using Clone Manager 8.0 (Sci-Ed Software) following the principles of AS RT-qPCR (Fig. 1a) and were synthesized by BGI (Beijing). For each mutation site, two forward primers were designed to specifically target the alleles of the wild type and the mutant, and the two alleles shared the same probe and the reverse primer. All mutation bases were situated at the 3’ end of the forward primers. In addition, an artificial mismatch close to the 3’ end was introduced into all the forward primers to enhance their discrimination performance. Parameters including hairpin formation, self-dimerization, and heterodimerization of all primer and probe sequences were analyzed using Clone Manager 8.0. The details of the primer-probe sets used for different assays, including sequences, modification, length, and product length, are summarized in Table S1. The method used for predictive value calculation for all mutation sites is provided in the Supporting Information.

2.2. Analytical performance of all AS RT-qPCR primers and probes

The analytical performance of the newly designed primer-probe sets was evaluated using individual RNA controls of Alpha, Beta, Gamma, Delta, and Kappa. However, for Lambda, Mu and Omicron, synthetic DNA positive controls were used since synthetic RNA was not yet commercially available. All RNA controls were quantified using RT-qPCR by N1 primer-probe set and standard curves generated using plasmids carrying the N1 gene. The concentrations of all DNA plasmids were determined using a Qubit dsDNA HS assay kit (Thermo Fisher, USA), and the copy number was calculated based on the length and Avogadro’s number. The details of RNA and DNA controls were provided in Table S2.

All standard curves were generated using a 10-fold serial dilution of the RNA or DNA for four gradients, with concentration levels ranging from approximately 10^4 to 10^5 copies/µL. The initial concentrations of all standard curves were reported in Table S2. All standard curves were tested in duplicate. Linear regression was performed using log copy number per µL and Ct values in RT-qPCR. The sensitivity was verified by the standard curves and the limits of detection (LODs). The LODs for all assays were verified using 10 replicates of positive controls with 10, 5, and 2.5 copies/µL to ensure >95% successful positive detection (Ct < 40).
The observed LODs were the lower one of 10, 5, and 2.5 copy/µL with >95% successful positive detection. And the estimated 95% LODs were defined by fitting a cumulative Gaussian distribution model (Ahmed et al., 2022) using GraphPad Prism 9.0. The specificity was demonstrated by detecting opposite genotypes to examine cross-reactivity of the two forward primers for a mutant and its wild type. The amplification efficiency of all assays was calculated by the following formula:

\[
\text{Amplification Efficiency (\%)} = 100 \times \left(1 - \frac{1}{\text{SLOPE}}\right)\]

2.3. One-step RT-qPCR conditions

TaqMan Fast Virus 1-step Master Mix (Thermo Fisher, USA) was employed for the one-step RT-qPCR process. The final concentrations of all primers and probes in each reaction were 500 nM and 250 nM, respectively. One-step RT-qPCR was performed on the Applied Biosystems™ QuantStudio™ 7 Flex Real-Time PCR system (Thermo Fisher, USA) with the following PCR cycler conditions: reverse transcription for 5 min at 50°C and initial denaturation for 20 s at 95°C, followed by 45 cycles of 5 s at 95°C and 30 s at 55°C. For every batch of RT-qPCR detection assays, a non-template control (NTC) was included.

2.4. Mock sample preparation

The viral signal in sewage may come from a mixture of multiple variants from different patients in the same catchment area. To simulate this feature, we prepared four different mixtures consisting of synthetic RNA controls of different SARS-CoV-2 variants, including one mixture containing equal volume (25 µL each) of four VOCs (Alpha, Beta, Gamma, and Delta) and three mixtures of two closely related variants (Delta and Kappa) at volume ratios of 1:4 (10 µL and 40 µL), 1:1 (25 µL each), and 4:1 (40 µL and 10 µL). The mock RNA samples from these mixtures were serially diluted tenfold in triplicate using DEPC-treated water to generate a gradient of four concentrations for VOCs mixture and a gradient of three concentrations for the mixture of Delta and Kappa. The initially spiked concentration of each variant RNA in each mixture was summarized in Figs. 5 and S4.

Quantification of four mixtures under each dilution concentration was conducted for all signature mutation sites that were carried by these variants using specific primer-probe sets, as well as the USCDC-recommended N1 targeted region. For each mock sample, the “detected concentration” of a mutation site was directly measured from the mock mixture, while the “expected concentration” was calculated by summing the values calculated from quantification results using individual synthetic RNA controls of different variants in that mixture.

2.5. Sewage sample collection, pre-treatment and RNA extraction

The E058 sample was collected from a manhole at 15 min intervals during 3 h in the morning peak (8 am to 11 am), and the quarantine hotel samples were taken at two periods (9 am to 12 am and 3 pm to 6 pm). All 30 mL of inactivated samples at 60°C for 30 min were firstly centrifuged at 4,750 g for 30 min, and then the supernatant was obtained for further ultracentrifugation at 150,000 g for 1 h (Allegra X-15R, Beckman Coulter, Indianapolis, IN) (Zheng et al., 2022). Finally, the concentrated pellet from the last step was re-suspended with ~200 µL PBS for RNA extraction. RNA was extracted using QIAamp Viral RNA Kits (Qiagen) according to the manufacturer’s instruction with a final elution volume of 50 µL.

2.6. Library preparation and sequencing

Targeted sequencing using the Nanopore platform for sewage samples was performed in this study. ARTIC amplicons (~400 bp) from
sewage samples were prepared following the ARTIC v3 protocol with some modifications, i.e., adding 5 µL of cDNA template instead of 2.5 µL. The purified amplicons were subjected to Nanopore sequencing by using an ONT ligation kit (SQK-LSK109) combined with an ONT Native Bar-coding Expansion kit (EXP-NBD104) following the manufacturer’s protocol. The library was run into a flow cell (FLO-MIN106) and sequenced on a GridION X5 device. The patient specimen sequence was obtained as previously described (Sit et al., 2020).

2.7. Raw data analysis

- The reads obtained from the Illumina sequencing for the patient specimen were aligned to the reference genome (MN908947.3) using bwa mem (0.7.17-r1198-dirty). Then, primer trimming, quality trimming (minimum quality of 20) and consensus calling (-m 2 -n A -a -d 0 -Q 0) were performed using iVar (1.3.1) and SAMtools mpileup (v1.13) following the iVar pipeline. The raw Nanopore reads were first base-called using Guppy (version 4.3.4) under the high-accuracy basecalling model, and then the consensus-level variant candidates were identified using the Medaka pipeline developed in the ARTIC workflow (https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html). The specific lineage of the called consensus sequence was identified by Nextclade (https://clades.nextstrain.org/). The mapping graph was generated by BLASTing the consensus sequences from sewage samples and patients in Geneious Prime 2021. The phylogenetic tree was created using Nextclade (https://clades.nextstrain.org/).

2.8. Statistical analysis

We used GraphPad Prism (GraphPad Software, San Diego, CA, USA), Microsoft Excel and the R package ggplot2 for linear regression and graph construction.

Supplemental method details, including RNA/DNA controls, matrix effect evaluation, sewage sample, and clinical information collection details, are provided in the Supporting Information.

3. Results

3.1. In silico performance evaluation of the mutation sites for variant identification

Different combinations from multiple representative mutation sites of the eight target variants were identified and evaluated for each variant (Fig. 1b). We observed that there was a wide range of specificity (42-100%) for variant detection based on a single mutation site (Fig. 2a and Table S3). Several mutation sites, specifically, K417N, K417T, 246-252 DEL, F490S, L452Q and E484Q, yielded ~100% specificity for the discrimination of the Beta, Gamma, Lambda and Kappa variant, respectively. However, using the mutation site N501Y alone, lower in silico specificities of 76% and 77% were obtained for the detection of Beta and Gamma, respectively, due to the interference of the large amount of Alpha variant genomes. As expected, in silico specificities were increased by combining two or three mutation sites for each variant. In particular, specificities of 97-100% for the Beta and Gamma variants could be achieved by using two- and three-site combinations (Fig. S2a).

In silico sensitivity of 85% to 100% was achieved when using only one single mutation site for each variant (Fig. 2b and Table S3). Specifically, the single mutation P681R had high in silico sensitivity for discriminating Delta and Kappa from other variants (99% and 95%, respectively).

Fig. 2. The predictive values of the combinations among signature mutation sites for the variants. a. Specificity values for the variant discrimination. b. Sensitivity values for the variant identification.
respectively). The *in silico* sensitivities decreased slightly when using combinations of two and three mutation sites, to 84-100% and 83-99%, respectively (Fig. S2b).

Those combinations that achieved both high specificity (>99%) and acceptable levels of sensitivity (>90%) for the target variants (Table 1) were selected for further evaluations. Specifically, Assay 1 for Alpha consists of N501Y and 69/70 DEL with 100% specificity and 97% sensitivity; Assay 2 for Beta consists of N501Y, E484K and K417N with 100% specificity and 92% sensitivity; Assay 3 for Gamma consists of N501Y, E484K and K417T with 100% specificity and 95% sensitivity; Assay 4 for Delta consists of T478K, P681R and L452R with 100% specificity and 99% sensitivity; Assay 5 for Omicron consists of N501Y, 69/70 DEL and T478K with 100% specificity and 98% sensitivity; Assay 7 for Mu consists of N501Y, E484K and P681H with 100% specificity and 98% sensitivity; and Assay 8 for Kappa consists of L452R, P681R and E484Q with 100% specificity and 98% sensitivity.

### 3.2. Analytical performance evaluation and sewage matrix effects

After selecting the above specific combinations of mutation sites for each variant based on *in silico* evaluations, twelve pairs of primer-probe sets targeting these different individual mutation sites were designed and assessed. All assays were able to detect the target variants even at lower concentrations of 5-10 viral copy/µL (Fig. 3) with amplification efficiencies above 85% (Table S4). The ranges of estimated LODs for all assays were from 2.6 to 8.9 copy/µL (Table S5). Notably, as shown in Fig. 4, *in cross-reactivity of the wild-type primer did not occur for the detection of Alpha, Beta, Delta, Omicron, Lambda, and Mu at concentrations up to 10^-3 viral RNA copy/reaction. For the assay of Kappa variant, cross-reactivity was not observed at concentrations up to 10^6 viral RNA copy/reaction. For Gamma, although cross-reactivity occurred for the K417T primer-probe set even at levels under 10^2 viral RNA copy/µL, this primer-probe set could still discriminate the mutant from the wild type at low concentrations due to the large difference in Ct values (>4) between the mutant and the wild type.

To investigate the effect of the sewage matrix on detection, the Ct values of the same amounts of SARS-CoV-2 variant RNA control in pure water and in RNA extracts from 30 mL negative sewage samples were compared. Technical replicates of each RNA control carrying target mutation sites were prepared and tested with the concentration levels of 10^1, 10^2, 10^3, and 10^4 viral RNA copy/µL. The sewage matrix did not significantly impact the detection of any of the four mutation sites across all the tested concentrations, resulting in a difference of less than 1 Ct from the values measured in pure water (Fig. S4 and Table S6). Moreover, for most of the assays, cross-reactivity was not observed when testing the SARS-CoV-2 variant RNA in sewage using wild-type primers.

### 3.3. Accuracy of signature mutation site quantification in mixtures of SARS-CoV-2 variants

A mixture of SARS-CoV-2 variants is likely present in a sewage sample, although the probability is relatively low when we use the sewage test for early warning purposes. The quantification accuracy of the designed primer-probe sets in the background of mixed variants was determined. Due to the negligible matrix effects in 30 mL sewage samples, the mock mixture samples were constructed using RNA standards in pure water.

For the mixture of four VOCs, eight mutation sites were tested and analyzed. As shown in Fig. 5a, all primer-probe sets correctly discriminated their target mutation sites in a mixture of variants, even for levels as low as 10 copy/µL (Fig. S5). The good linear regression for the dilution series (R^2 > 0.9) revealed the reliability of the quantification method even at low levels of target mutation sites with negligible cross-reactivity to the corresponding wild-type genotypes. In addition, the concentrations of all target mutation sites were accurately quantified, with 7 out of 9 mutation sites exhibiting less than 10-fold deviation from the expected values. The exceptions were L452R and 69/70 DEL, which had large deviations (20-40 fold) from the expected values. In addition, when compared to N1, all primer-probe sets had comparable levels of sensitivity and quantification accuracy.

For three mixtures containing Delta and Kappa at three different ratios, as shown in Figs. 5 and S5, all five pairs of primer-probe sets were sensitive and reliable even at 10 copy/µL for the detection and quantification of five mutation sites, including three common sites (Loci 452, 681, and N1) and two characteristic sites, E484Q for Kappa and T478K for Delta. The good linear regression for the dilution series (R^2 > 0.9) suggested negligible cross-reaction effects, and the concentrations detected by all assays were within 10-fold deviation from the expected values. Additionally, the presence of Delta and Kappa in the three scenarios was well illustrated by combining shared mutation sites and unique signature assays of T478K and E484Q.

### 3.4. Implementation of the novel method

#### 3.4.1. Practical effectiveness of assays for Beta variant identification

One implementation case was the integration of the novel method into routine sewage surveillance in Hong Kong, which covers more than 85% of the total population. The detection method employed for routine sewage testing was based on RT-qPCR using the USCDC-recommended N1 assay. On 2 May 2021, a sewage sample E058 collected from a site covering 58,000 residents was tested positive for SARS-CoV-2, yielding a concentration of 41,000 copy/L. On 7 May 2021, a patient who lived within the sewershed of the positive sampling site was diagnosed with Beta variant infection. No other COVID-19 patient was found in that sewershed 14 days before or after sewage sampling, indicating that the patient was the sole contributing source to the positive signal of SARS-CoV-2 picked up by sewage surveillance. Applying our method to this positive sewage sample identified the presence of Beta as indicated by the concurrent positive signals of N501Y and E484K and simultaneous absence of the corresponding wild-type genotypes (i.e., N501 and E484), in addition to the positive result for K417N and negative result for K417T. Comparisons of viral concentrations detected by different primer-probe sets revealed that although the Ct values obtained by primer-probe sets targeted at mutation sites were 3-6 larger than that of the N1 assay for the same sewage sample, the viral copy numbers calculated using their standard curves were generally consistent with

### Table 1

| Assay | Composition | Variant | In silico sensitivity (%) | In silico specificity (%) |
|-------|-------------|---------|--------------------------|--------------------------|
| Assay-N501Y Alpha | 1 | 69/70 DEL | 97 (898,890/924,996) | −100 (3,404,043/3,404,196) |
| Assay-N501Y Beta | 2 | E484K K417N | 92 (22,350/24,192) | −100 (4,302,611/4,304,400) |
| Assay-N501Y Gamma | 3 | L452R P681R | 95 (82,790/87,226) | −100 (4,241,353/4,241,366) |
| Assay-L452R Delta | 4 | T478K | 99 (2,424,367/2,438,114) | −100 (1,885,908/1,890,478) |
| Assay-N501Y Omicron | 5 | L452Q Lambda | 90 (1,150,1,284) | −100 (4,327,251/4,327,308) |
| Assay-L452Q Lambda | 6 | F490S | 98 (8,742/8,945) | −100 (3,419,614/3,419,647) |
| Assay-N501Y Mu | 7 | E484K P681H | 98 (6,527/6,634) | −100 (4,314,054/4,321,958) |
| Assay-L452R Kappa | 8 | P681R E484Q | 98 (4,921,5,012) | −100 (4,318,356/4,323,580) |
each other, ranging from 20,992 to 53,765 copy/L (Table S7).

We also acquired the partial sequences of the virus via Nanopore sequencing. Although the partial sequences could confirm that it was the Beta variant (Fig. S6), the completeness of the consensus sequence was only 66%, and the characteristic mutation sites of N501Y, K417N and E484K were not covered by the sequences, perhaps due to fragmentation of viral RNA in sewage.

3.4.2. Identification of Delta and Omicron variants in quarantine hotels

Another implementation case was the detection of Delta importation at a quarantine hotel. As of November 2021, there has been no Delta outbreak in the local community in Hong Kong. Designated hotels for compulsory quarantine of inbound travelers were hotspots for the importation of COVID-19 cases. Sewage samples were consecutively collected from a local quarantine hotel for two weeks of continuous monitoring in September 2021. The Delta variant, having three signature mutation sites of L452R, T478K, and P681R, was detected in three positive sewage samples, which had low (1,894 copy/L), medium (8,352 copy/L), and high (606,516 copy/L) viral concentrations. The epidemiological data agreed with the sewage testing results, revealing that a carrier of the Delta variant stayed in the hotel during the sampling period and contributed exclusively to the positive signals in these three samples (Fig. 6 a). The sewage samples were tested negative for SARS-CoV-2 after this carrier was moved to a hospital for treatment.

In addition, the newly emerging Omicron variant has become a greater threat since the end of November 2021. Positive results for three signature mutation sites of N501Y, 69/70 DEL and T478K, as well as a negative result for L452R, were detected in three sewage samples (RGH-01, RGH-02 and RGH-03) collected from another quarantine hotel on 15 and 16 December 2021, indicating the presence of the Omicron variant. Later, a patient was diagnosed on 18 December 2021 and confirmed to carry the Omicron variant on 19 December, which agreed with the sewage testing results.

We compared the concentrations of the variants in sewage samples measured by our variant assays with those measured by the USCDC N1 assay (Fig. S7). The results revealed that viral concentrations measured by three characteristic mutation sites of Delta in the same sewage sample were very close to each other, with differences of Ct values less than 2. Although a large Ct difference was observed with the N1 assay (2-7 cycles higher on our assay than on N1), the viral concentration calculated with the standard curve of N1 yielded similar viral concentrations, within a 4-fold range. For the detection of Omicron, comparable results were obtained between the 69/70 DEL and N1 assay. However, the concentrations calculated from N501Y, T478K and L452 were lower than those calculated from the N1 assay. Moreover, for some samples with the lower viral titers, a few assays, such as T478K, even reported concentrations below the quantification limits.

We further compared the genome sequences of the Delta variant obtained by the Nanopore sequencing platform from three positive sewage samples to those obtained by the Illumina platform from swab samples.
specimens from the matched patient (Fig. 6). Nanopore sequencing generated near-complete (>98.9%) genome sequences from all three sewage samples with average sequencing coverages of 467, 161 and 267, respectively. Phylogenetic analysis showed that the four consensus genome sequences determined from three sewage samples and the patient specimen were closely related to each other within the Delta lineage (Fig. S6). In addition, we found that the consensus genome sequence from the sewage sample with the highest viral concentration achieved the best concordance with that from the matched patient, with only four undetected substitutions in the sewage sample. For the other two sewage samples, which had medium and low viral concentrations, there were 10 and 15 undetected substitutions, respectively (Table S8).

Fig. 4. Analytical performance of assays for the mutation detection. Solid lines indicate the standard curves for the mutant assays. Dash lines indicate the specificity of assays for the wild type in detecting the mutant. Colors imply different primer-probe sets. All details of all standard curves were provided in Table S4.
caused by fragmentation of the viral RNA or by the lower viral load in the sewage sample.

4. Discussion

Compared with viral whole-genome sequencing, PCR-based technologies such as RT-qPCR, nested RT-PCR, and AS RT-qPCR represent rapid, sensitive, cost-effective assessment tools for extensive screening of SARS-CoV-2 variants in sewage surveillance. The AS RT-qPCR approach, which simultaneously measures multiple mutant and wild-type sites, provides a higher level of confidence in discriminating variants. In this study, novel real-time RT-qPCR assays for allelic discrimination were developed and evaluated for the detection of eight SARS-CoV-2 variants based on 12 mutation sites in the spike protein gene of the virus. A previously reported method focused on the use of three mutation sites for detection of Alpha (Lee et al., 2021). Our study represents an expansion of the previously published AS RT-qPCR methods for SARS-CoV-2 variant detection, and the combinations of signature mutation sites designed for specific variants can be quickly pivoted for newly emerging variants, as demonstrated in the application of Omicron variant detection in this study.

The selection of signature mutation sites for detecting different SARS-CoV-2 variants is the first step toward obtaining conclusive results in variant discrimination. In this study, the discriminatory performance of the multiple mutation sites located in the spike protein gene of the virus was thoroughly assessed using the SARS-CoV-2 genome database of GISAID as of 9 December 2021. The spike protein is of clinical importance as a target for vaccine development and antigen testing (Harvey et al., 2021; Li et al., 2020). Although in silico analysis using the current database of the virus genomes may show that the selection of a single mutation site can have nearly 100% specificity for a specific variant, the apparent high specificity could be biased by low numbers of a target variant and large numbers of nontarget sequences of other variants in the current database. In addition, the fragmented viral genomes present in sewage samples (Graber et al., 2021) make detection using a single mutation site prone to false negative results. Thus, in our study, a combination of two or three signature mutation sites was designed for each target SARS-CoV-2 variant. The designed assays for all variants achieved high sensitivity (90-100%) and specificity (~100%). Assay sensitivity and specificity are crucial for identifying and distinguishing specific variants. All assays could detect target variants at lower concentrations with the estimated LODs ranging from 2.6 to 8.9 viral RNA copy/µL, and the inhibitory matrix effects of 30 mL sewage were negligible for all mutant assays. Cross-reactivity between opposite genotypes was minimal, and the method reliably quantified target mutation sites even when multiple SARS-CoV-2 variants were concurrently present in the sewage. Although assays for a few mutation sites, such as K417T, had cross-reactivity even below 100 copy/µL, they were still able

![Graph A](image1.png)

**Fig. 5.** The detected value and the expected value of mutation sites in the mock RNA mixtures. a. VOCs mixture detected in first dilution concentration. Mixtures of Delta and Kappa in volume ratio of 1:1 (b), 4:1 (c) and 1:4 (d) for the first dilution concentration. The dash lines respectively indicated the detected value equal to the expected value and the detected value within 10-fold to the expected value. The initial spiked concentration of each variant in each reaction quantified by N1 primer-probe set was shown in parentheses. The circles indicate mutant assays, and the stars indicate the wild-type assays.
to be discriminated from the opposite wild-type genotype, due to the higher sensitivity for the mutant assay and the large differences in Ct values between the mutant assay and the wild-type assay. The variant can be identified by diluting the sample to a concentration that is low enough to be recognized by the mutant assay but not by the wild-type assay. Besides, the variant type can also be discriminated by combining the results of multiple targets of a variant.

Rapid variant detection in sewage samples can help quash an outbreak before it spreads. Our method was incorporated into the established routine sewage surveillance in Hong Kong and strengthened public health surveillance by providing real-time determination of variant types to allow assessment of the associated risks, demonstrating the application feasibility. Moreover, considering the present epidemiological trends, a practical workflow for variant detection was developed for prompt public health decisions (Fig. 1c). In this workflow, a hierarchical detection framework according to the current variants and shared mutation sites among different variants was proposed. The mutation sites of N501Y and L452R were selected as the first layer of screening, since they are shared by multiple variants; L452Q can be used as an alternative to identify the Lambda variant. Then, three shared mutation sites were selected to further discriminate specific variants at the second layer. In the third layer, the variant type was finally identified according to the unique signature mutation sites. Notably, for the Alpha variant, two mutation sites, N501Y and 69/70 DEL, are shared with the recently emerging Omicron variant; therefore, A570D was further selected as a unique signature mutation site with an in silico specificity of 99% and an in silico sensitivity of 98%. We used primer-probe sequences designed by another group (Lee et al., 2021) to identify A570D and validated that the assay had sufficient analytical sensitivity and differentiable specificity using an Alpha RNA control (Fig. S8). Additionally, due to the uncertainty of the sewage samples, if no signal is detected for either the wild type or the mutant type in the first two layers, the signature mutation sites in the next layer could be applied according to the prevalence trend. This workflow was designed based on the current epidemiological trends, and it should be modified and updated with new developments and as new variants emerge. In addition, for one variant, assays consisting of three primer-probe sets could be streamlined to two sets by excluding those assays at the second layer according to the specific situation, providing flexibility in the application (Table S9). Furthermore, the Assay 5 designed for the Omicron BA.1 will not work for the Omicron BA.2 that does not carry 69/70 DEL. However, the streamlined Assay 5 with the combination of N501Y and T478K is applicable for the detection of Omicron BA.2. Specifically, two descendants of Omicron BA.2, the currently circulating BA.2.12.1 worldwide and the BA.2.2 prevalent in the fifth wave outbreak in Hong Kong, can be distinguished according to the presence or absence of the L452Q mutation site. It is worth noticed that for the two newest emerging Omicron sublineages, BA.4 and BA.5, the identification assays need to be reconfigured. For example, the new combinations of L452R, N501Y, and 69/70 DEL could be used as signature mutation sites to detect these two variants.

It should be noted that while the proposed AS RT-qPCR method can identify different SARS-CoV-2 variants with very high probability, viral whole-genome sequencing is suggested to be conducted for the final confirmation of the variant type. However, the sequencing-based approach requires a turnaround time of a few days (Chiara et al., 2021) and may be more uncertain due to the low variant abundance in sewage, as demonstrated in this study by the examples of Beta and Omicron variants with incomplete consensus sequences.

5. Conclusion

Within this study, we developed eight assays to robustly identify eight variants. Firstly, a set of signature mutation sites were identified based on in silico sensitivity and specificity for five variants of concern (VOCs), two variants of interest (VOIs), and one variant under monitoring (VUM). Besides, novel primer-probe sets with high sensitivity were designed for identification of 12 mutation sites at the spike gene of the SARS-CoV-2 genome, and eight novel AS RT-qPCR assays using different combinations of these 12 mutation sites to distinguish eight SARS-CoV-2 variants were developed, which can achieve ~100% specificity and ~90% sensitivity to identify eight SARS-CoV-2 variants in sewage samples using their unique assays. Furthermore, practical implementation effectiveness of the method was proven by accurately identifying the Beta, Delta and Omicron variants in sewage samples prior to the clinical variant confirmation of COVID-19 patients in Hong Kong. Finally, a hierarchical detection workflow was developed for flexible deployment of the method for eight individual SARS-CoV-2
variants and other potential newly emerging variants. In conclusion, our AS RT-qPCR variant detection assays enable sensitive, extensive, and real-time surveillance for SARS-CoV-2 variants in sewage. Besides, the design strategy could be employed to novel emerging variants. These assays could be integrated into the existing community-level COVID-19 sewage surveillance systems to provide early warning signals of potential outbreaks to boost preparedness and take public health actions based on high-confidence identification of variants in sewage.

**Funding**

This study was financially supported by Health and Medical Research Fund (HMRF) (COVID193015) and (COVID190205), the Food and Health Bureau, the Government of the Hong Kong Special Administrative Region (SAR), China.

**Data availability**

The consensus sequences of the patient specimen and sewage samples were available in the NCBI under the submission number SUB11224524.

**CRediT authorship contribution statement**

Xiaoxing Xu: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. Yu Deng: Methodology, Formal analysis, Data curation, Writing – review & editing. Jiahui Ding: Methodology, Investigation. Xiawan Zheng: Investigation. Shuxian Li: Investigation. Lei Liu: Methodology. Ho-kwong Chui: Resources, Supervision, Writing – original draft, Writing – review & editing. Tong Zhang: Conceptualization, Methodology, Resources, Supervision, Writing – original draft, Writing – review & editing.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgements**

We thank the Environmental Protection Department (EPD) and Drainage Services Department (DSD) of Hong Kong SAR Government for sewage sample collection. Xiaoxing Xu, Jiahui Ding, Xiawan Zheng and Shuxian Li would like to thank for the University of Hong Kong for the Postgraduate Studentship (PGS). Technical support from Ms. Vicky Fung is greatly appreciated.

**Supplementary materials**

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2022.118686.

**References**

Agrawal, S., Orschler, I., Schubert, S., Zachmann, K., Heijnen, L., Tavazzi, S., Gawlik, B. M., de Graaf, M., Medema, G., Lackner, S. 2022. Prevalence and circulation patterns of SARS-CoV-2 variants in European sewage mirror clinical data of 54 European cities. Water Res. 214, 118162.

Ahmed, W., Angel, N., Edson, J., Bibby, K., Orschler, L., Schubert, S., Zachmann, K., Heijnen, L., Tavazzi, S., Gawlik, B., Poon, L.L., Zhang, X., Alain, T. 2021. Quantitative analysis of SARS-CoV-2 RNA from wastewater solids in communities with low COVID-19 incidence and prevalence. Water Res. 188, 116566.

Deng, Y., Xu, Z., Zheng, X., Ding, J., Li, S., Chui, H.-K., Wang, T.-K., Poon, L.L., Zhang, T. 2022a. Use of sewage surveillance for COVID-19 to guide public health response: A case study in Hong Kong. Sci. Total Environ. 153233.

Deng, Y., Zheng, X., Xu, Z., Chui, H.K., Lai, W.K., Li, S., Tun, H.M., Poon, L.L., Ding, J., Peiris, M., Leung, G.M., Zhang, T. 2022b. Use of sewage surveillance for COVID-19: a large-scale evidence-based program in Hong Kong. Environ. Health. Perspect. 130 (5), 057008 https://doi.org/10.1289/EHP966.

Gaebler, C., Wang, Z., Lorenzi, J.C., Mueckf, F., Finkin, S., Tokuyama, M., Cho, A., Jankovic, M., Schafer-Babajee, D., Oliveira, T.Y. 2021. Evaluation of antibody immunity to SARS-CoV-2. Nature 591 (7851), 639-644.

Gazdar, F.M., Nuhmananda, M., Nabhila, R., Supertiyti, E., Hakim, M.S., Arguni, E., Daniswiyana, E.W., Nuryastuti, T., Haryama, S.M., Bihawa, T. 2021. Detection of SARS-CoV-2 spike protein D614G mutation by qPCR-hRM analysis. Helinyon 7 (9), 07936.

Gibas, C., Lambirh, K., Mittal, N., Jou, M.A.L., Barua, V.B., Brazell, L.R., Hinton, K., Lontai, J., Stark, N., Young, I. 2021. Implementing building-level SARS-CoV-2 wastewater surveillance on a university campus. Sci. Total Environ. 782, 146749.

Grabar, T.E., Mercier, E., Bhagmar, K., Fuzzen, M., D’Aoust, P.M., Huang, H.-D., Tian, X., Towhid, S.T., Plaza-Diaz, J., Eid, W. 2021. Near real-time determination of B. 1.1. 7 in proportion to total SARS-CoV-2 viral load in wastewater using an allele-specific primer extension PCR strategy. Water Res. 205, 117681.

Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., Peacek, S.J. 2021. SARS-CoV-2 variants, spike mutations and immune escape. Nat. Rev. Microbiol. 19 (7), 495-424.

Izquierdo-Lara, R., Etlinga, G., Heijnen, L., Mumma, B.B.O., Schapendonk, C.M., Nieuwenhuijse, D., Kon, M., Lu, L., Aarestrup, F.M., Lycett, S. 2021. Monitoring SARS-CoV-2 circulation and diversity through community wastewater sampling, the Netherlands and Belgium. Emerg. Infect. Dis. 27 (5), 1405.

Kalderis, T., Dabbas, T., Jerzem, M., Assa, S., Charara, A. 2022. Modified “Allele-specific qPCR” method for SNP genotyping by FRET. Front. Plant Sci. 12.

Keshaviah, A., Hu, X.C., Henry, M. 2021. Developing a flexible national wastewater surveillance system for COVID-19 and beyond. Environ. Health. Perspect. 129 (4), 045002.

Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E.E., Bhattacharya, T., Foley, B. 2020. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 182 (4), 812-827 e819.

Korfmaner, K.S., Harris-Lovett, S., Nelson, K.L. 2021. Campus collaborations as a model for transforming SARS-CoV-2 wastewater surveillance research into public health action. Environ. Sci. Technol.

Lee, W.L., Imakawa, M., Armas, F., McElroy, K.A., Gu, X., Duvallet, C., Chandra, F., Chen, H., Leifeld, M., Mendola, S. 2021. Quantitative SARS-CoV-2 alpha variant B.1.1.7 tracking in wastewater by allele-specific RT-qPCR. Environ. Sci. Technol. Lett. 8 (8), 675-682.

Lefèvre, S., Rihan, A., Van der Meulen, J., Pattyn, F., Van Maerken, T., Van Dorpe, J., Alland, D., Bivins, A., Bivins, K. 2021. Cost-effective and robust genotyping using double-mismatch allele-specific quantitative PCR. Sci. Rep. 9 (1), 1-11.

Li, Q., Wu, J., Nie, J., Zhang, L., Hao, H., Liu, S., Zhao, C., Zhang, Q., Liu, H., Nie, L. 2020. The impact of mutations in SARS-CoV-2 spike on viral infectivity and cell entry. Cell 182 (5), 1248-1294 e1289.

Mallapaty, S. 2020. How sewage could reveal true scale of coronavirus outbreak. Nature 580 (7802), 176-177.

Mayne, J., Zhang, X., Alain, T. 2021. Impact of SARS-CoV-2 spike mutations on viral infectivity. J. Mol. Biol. 432 (19), 5212-5226.

Matsubara, T., Iwata, D., Wilton, M., Zambon, M., Bentley, E., Buikaji, E., Fritzschke, M., Rake, M., Majumdar, M. 2020. Tracking SARS-CoV-2 in sewage: evidence of changes in virus variant predominance during COVID-19 pandemic. Viruses 12 (10), 1144.
Medema, G., Been, F., Heijnen, L., Petterson, S., 2020. Implementation of environmental surveillance for SARS-CoV-2 virus to support public health decisions: Opportunities and challenges. Curr. Opin. Environ. Sci. Health 17, 49–71.

O’Meara, D., Ahmadian, A., Odberg, J., Lundeberg, J., 2002. SNP typing by apyrase-mediated allele-specific primer extension on DNA microarrays. Nucleic Acids Res. 30 (15), e75.

Parunu, D.V., 2021. Revised World Health Organization (WHO) terminology for variants of concern and variants of interest of SARS-CoV-2. Med. Sci. Monit. 27, e933622.

Peccia, J., Zallii, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-Massana, A., Ko, A.I., Malik, A.A., Wang, D., Wang, M., 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. Nat. Biotechnol. 38 (10), 1164–1167.

Peterson, S.W., Lidder, R., Daigle, J., Wonitowy, Q., Dueck, C., Nagasawa, A., Mulvey, M. R., Mangat, C.S., 2022. RT-qPCR detection of SARS-CoV-2 mutations S 69–70 del, S N501Y and N D3L associated with variants of concern in Canadian wastewater samples. Sci. Total Environ. 810, 151283.

Sit, T.H., Brackman, C.J., Ip, S.M., Tam, K.W., Law, P.Y., To, E.M., Veronica, Y., Sims, L. D., Tsang, D.N., Chu, D.K., 2020. Infection of dogs with SARS-CoV-2. Nature 586 (7831), 776–778.

Team, E.E., 2021. Updated rapid risk assessment from ECDC on the risk related to the spread of new SARS-CoV-2 variants of concern in the EU/EEA--first update. Eurosurveillance 26 (3), 2101211.

Vogels, C.B., Breban, M.I., Ott, I.M., Alpert, M.E., Petrone, A.E., Kalinich, C. C., Earnest, R., Rothman, J.E., Goes de Jesus, J., 2021. Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. PloS Biol. 19 (5), e3001236.

Wang, Z., Schmidt, F., Weisblum, Y., Muecksh, F., Barnes, C.O., Finkin, S., Schafer-Babajew, D., Cipolla, M., Goebrer, C., Lieberman, J.A., 2021. mRNA vaccine-elicted antibodies to SARS-CoV-2 and circulating variants. Nature 592 (7855), 616–622.

Wu, D.Y., Ugozzioli, L., Pal, B.K., Wallace, R.B., 1989. Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia. Proc. Natl. Acad. Sci. 86 (8), 2757–2760.

Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bonneau, R., Brown, M.A., Bushman, M., Chai, P.R., 2022. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamic and clinical presentation of new COVID-19 cases. Sci. Total Environ. 805, 150121.

Xu, X., Deng, Y., Zheng, X., Li, S., Ding, J., Yang, Y., On, H.Y., Yang, R., Chui, H.K., Yau, C.L., Tun, H.M., Chin, A.W.H., Poon, L.L., Peiris, M., Leung, G.M., Zheng, T., 2022. Evaluation of RT-qPCR primer-probe sets to inform public health interventions based on COVID-19 sewage tests. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.2c00974. In press.

Xu, X., Zheng, X., Li, S., Lam, N.S., Wang, Y., Chu, D.K., Poon, L.L., Tun, H.M., Peiris, M., Deng, Y., 2021. The first case study of wastewater-based epidemiology of COVID-19 in Hong Kong. Sci. Total Environ., 148000.

Yaniv, K., Ozer, E., Lewis, Y., Kushnarov, A., 2021. RT-qPCR assays for SARS-CoV-2 variants of concern in wastewater reveals compromised vaccination-induced immunity. Water Res. 207, 117808.

Zhang, T., 2022. Wastewater as an information source of COVID-19. Sci. Bull. https://doi.org/10.1016/j.scib.2022.04.006, In press.

Zheng, X., Deng, Y., Xu, X., Li, S., Zhang, Y., Ding, J., On, H.Y., Lai, J.C., Yau, C.L., Chiu, A.W., 2022. Comparison of virus concentration methods and RNA extraction methods for SARS-CoV-2 wastewater surveillance. Sci. Total Environ. 824, 153687.