

# Inferring transmission fitness advantage of SARS-CoV-2 variants of concern from wastewater samples using digital PCR, Switzerland, December 2020 through March 2021

Lea Caduff<sup>1</sup>, David Dreifuss<sup>2,3</sup>, Tobias Schindler<sup>4,5</sup>, Alexander J Devaux<sup>1</sup>, Pravin Ganesanandamoorthy<sup>1</sup>, Anina Kull<sup>1</sup>, Elyse Stachler<sup>1</sup>, Xavier Fernandez-Cassi<sup>6</sup>, Niko Beerenwinkel<sup>2,3</sup>, Tamar Kohn<sup>6</sup>, Christoph Ort<sup>1</sup>, Timothy R Julian<sup>1,4,5</sup>

1. Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

2. Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

3. SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

4. Swiss Tropical and Public Health Institute, Basel, Switzerland

5. University of Basel, Basel, Switzerland

6. Laboratory of Environmental Chemistry, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Correspondence: Timothy R. Julian (tim.julian@eawag.ch)

## Citation style for this article:

Caduff Lea, Dreifuss David, Schindler Tobias, Devaux Alexander J, Ganesanandamoorthy Pravin, Kull Anina, Stachler Elyse, Fernandez-Cassi Xavier, Beerenwinkel Niko, Kohn Tamar, Ort Christoph, Julian Timothy R. Inferring transmission fitness advantage of SARS-CoV-2 variants of concern from wastewater samples using digital PCR, Switzerland, December 2020 through March 2021. *Euro Surveill.* 2022;27(10):pii=2100806. <https://doi.org/10.2807/1560-7917.ES.2022.27.10.2100806>

Article submitted on 02 Aug 2021 / accepted on 21 Jan 2022 / published on 10 Mar 2022

**Background:** Throughout the COVID-19 pandemic, SARS-CoV-2 genetic variants of concern (VOCs) have repeatedly and independently arisen. VOCs are characterised by increased transmissibility, increased virulence or reduced neutralisation by antibodies obtained from prior infection or vaccination. Tracking the introduction and transmission of VOCs relies on sequencing, typically whole genome sequencing of clinical samples. Wastewater surveillance is increasingly used to track the introduction and spread of SARS-CoV-2 variants through sequencing approaches. **Aim:** Here, we adapt and apply a rapid, high-throughput method for detection and quantification of the relative frequency of two deletions characteristic of the Alpha, Beta, and Gamma VOCs in wastewater. **Methods:** We developed drop-off RT-dPCR assays and an associated statistical approach implemented in the R package WWdPCR to analyse temporal dynamics of SARS-CoV-2 signature mutations (spike  $\Delta 69-70$  and ORF1a  $\Delta 3675-3677$ ) in wastewater and quantify transmission fitness advantage of the Alpha VOC. **Results:** Based on analysis of Zurich wastewater samples, the estimated transmission fitness advantage of SARS-CoV-2 Alpha based on the spike  $\Delta 69-70$  was 0.34 (95% confidence interval (CI): 0.30–0.39) and based on ORF1a  $\Delta 3675-3677$  was 0.53 (95% CI: 0.49–0.57), aligning with the transmission fitness advantage of Alpha estimated by clinical sample sequencing in the surrounding canton of 0.49 (95% CI: 0.38–0.61). **Conclusion:** Digital PCR assays targeting signature mutations in wastewater offer near real-time monitoring of SARS-CoV-2 VOCs and potentially earlier detection and inference on transmission fitness advantage than clinical sequencing.

## Introduction

Emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genetic variants of concern (VOCs) can impact negatively on efforts to reduce the global disease burden attributable to coronavirus disease (COVID-19). New VOCs emerge when mutations in the viral genome increase transmissibility and/or virulence or reduce neutralisation by antibodies obtained from prior infection or vaccination. Notable examples include Alpha (Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin) designation lineage B.1.1.7, first described in the United Kingdom in September 2020), Beta (lineage B.1.351, first discovered in South Africa in December 2020) and Gamma (lineage P.1, first discovered in Brazil in January 2021).

During the timeline of this study (December 2020 through March 2021), Alpha, Beta and Gamma variants spread worldwide, with Alpha overtaking other variants to become the most abundant variant in Europe. The rise in the Alpha variant coincided with a rise in the number of reported infections, motivating Alpha as the target VOC for this study. The VOCs are identified by combinations of co-occurring mutations. Tracking the introduction and transmission of VOCs relies on sequencing, typically whole genome sequencing of clinical samples. Given the need for fast and cheap methods that allow high throughput and are easy to interpret, PCR-based methods targeting signature or defining mutations (typically point mutations or deletions) have become increasingly common [1,2].

**TABLE 1**

Primer and probe sequences for SARS-CoV-2 drop-off RT-dPCR, including resulting amplicon size and corresponding reference

DNA oligonucleotide name	DNA sequence (5'–3')	Reference
<b>Target: spike Δ69–70 (108 bp amplicon)</b>		
Yale_69–70_F	TCAACTCAGGACTTGTCTTACCT	[1]
Yale_69–70_R	TGGTAGGACAGGGTTATCAAAC	[1]
<i>Deletion probe</i>		
Yale_69–70_Cy5_P	Cy5-TTCCATGCTATACATGTCTCTGGGA-BHQ2	[1], new quencher, fluorophore
<i>Universal probe</i>		
LC_69–70_HEX_P	HEX-CCAATGGTACTAAGAG-MGBQ530	This study
<b>Target: ORF1a Δ3675–3677 (128 bp amplicon)</b>		
Yale_ORF1a-del_F	TGCCTGCTAGTTGGGTGATG	[1]
Yale_ORF1a-del_R	TGCTGTCATAAGGATTAGTAACACT	[1]
<i>Deletion probe</i>		
Yale_ORF1a-HEX_P	HEX-GTTTGTCTGGTTTAAGCTAAAAGACTGTG-BHQ1	[1], new quencher, fluorophore
<i>Universal probe</i>		
LC_ORF1a-Cy5_P	Cy5-CGTATTATGACATGGTTGGATATGGTTGAT-BHQ2	This study

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Because SARS-CoV-2 RNA is shed in faeces and other bodily fluids [3,4], it is readily detectable in wastewater samples. Surveillance for SARS-CoV-2 in wastewater (wastewater-based epidemiology) is used to inform COVID-19 disease trajectories in catchment areas. Wastewater can also be used to inform the introduction and spread of SARS-CoV-2 variants. To date, sequencing SARS-CoV-2 in wastewater has been applied to detect novel variants in the San Francisco Bay Area [5], detect and track geographical spread of the Alpha variant in Switzerland [6] and track the increased occurrence of Alpha in London [7]. However, in contrast to clinical samples, wastewater samples consist of a mix of SARS-CoV-2 variants representative of the variants circulating in the catchment areas. Therefore, VOCs in wastewater are inferred by tracking combinations of multiple signature or characteristic mutations. Confidence in the detection can be increased by investigating temporal trends and/or tracking signature mutations that co-occur on a single amplicon [6].

Although sequencing SARS-CoV-2 RNA in wastewater is a promising approach to track VOC from hundreds or thousands of individuals in a single sample, there are temporal lags in sample collection and analysis, similar to lags associated with clinical testing. Sequencing and bioinformatics pipelines typically take at least 3 days to yield results after sample RNA is collected. For more rapid identification of VOCs, PCR-based assays have been developed to identify their characteristic mutations [1,2,8]. Examples include assays targeting single-nucleotide polymorphisms, such as N501Y and E484K in the spike protein, or characteristic deletions such as spike Δ69–70 and ORF1a Δ3675–3677. The PCR-based assays, designed for detection of variants

in clinical samples, have also been applied successfully for detection in wastewater [9–11]. Using quantitative PCR-based assays (RT-qPCR or RT-dPCR), the proportion of a specific variant within the virus population can be inferred. Proportions are determined using quantitative assays by comparing the quantity of the target gene region characteristic of the variant to the quantity of a conserved region either on a distinct gene (i.e. N gene) or on a region adjacent to the VOC-specific region [9–11]. Although this approach has demonstrated correlation with proportions of variants in clinical samples, bias is introduced by variation in efficiency and sensitivity of the quantification approaches of two different assays.

Here, we developed and assessed methods for inferring the epidemiology of SARS-CoV-2 VOCs from wastewater using digital RT-dPCR assays with non-competitive hydrolysis probes targeting signature mutations of the Alpha, Beta and Gamma VOCs. We aimed to demonstrate that the approach can be used to quantify transmission fitness advantage and provide data similar to those obtained from clinical samples. Overall, our aim was to demonstrate the utility of digital PCR assays targeting signature mutations in wastewater in providing near real-time monitoring of SARS-CoV-2 VOCs and potentially earlier inference on the epidemiology of VOC transmission relative to inferences based on clinical samples.

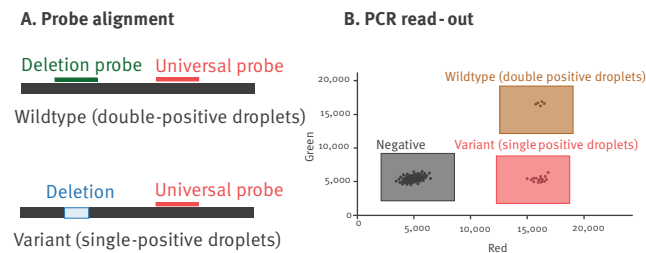
## Methods

### Overview

We adapted and applied a rapid, high-throughput method for detection and quantification of the

## FIGURE 1

### Schematic overview of the SARS-CoV-2 drop-off RT-dPCR assay based on two different probes, a deletion probe and a universal probe



SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Panel A: The universal probe binds to both the wildtype and the variant, while the deletion probe only binds to the wildtype (no deletion). Panel B: The wildtype appears as double-positive droplets (brown) and the variant as single-positive droplets (red) in the associated dPCR 2D scatter plot.

frequency of two signature deletion mutations characteristic of the Alpha, Beta and Gamma VOCs in wastewater. We quantify proportions of the variant in a wastewater sample using two non-competitive hydrolysis probes targeting conserved regions on the same amplicon within a single RT-dPCR assay (drop-off RT-dPCR assay). This allowed for simultaneous quantification of wildtype (strain not carrying the mutation of interest) and variants. We applied the assay to raw influent obtained from a wastewater treatment plant in Zurich, Switzerland, over the period December 2020 to March 2021 when Alpha was introduced and spread throughout the region [12]. We further developed a statistical approach to analyse temporal dynamics in drop-off RT-dPCR data to quantify transmission fitness advantage, and compared the results to those obtained from analysis of clinical samples.

### Drop-off RT-dPCR assays for variant of concern detection in wastewater

We developed a first drop-off RT-dPCR assay using a hydrolysis probe to target the spike  $\Delta 69-70$  deletion characteristic of the SARS-CoV-2 Alpha variant, and a second assay for ORF1a  $\Delta 3675-3,677$  characteristic of the Alpha, Beta, and Gamma variants. Although these deletions are characteristic of VOCs, both have been detected in other variants, including non-VOCs [1]. Notably, spike  $\Delta 69-70$  is associated with increased cell infectivity: in the Alpha variant, it compensates for reduced infectivity caused by immune escape mutations [13,14]. The assay adapts the approach of Vogels et al. in which the hydrolysis probes anneal preferably on amplicons without the deletion and are hereafter referred to as ‘deletion probe’ [1]. Based on RT-qPCR analysis of a suite of 71 previously sequenced clinical samples, Vogels et al. report that the hydrolysis probes only anneal to the samples without the deletion [1].

To measure the fraction of VOCs, we adapted the assay by designing additional hydrolysis probes which target conserved regions on the same amplicon as the deletion probes (Table 1). These probes, hereafter referred to as ‘universal probes’, detect amplicons of all lineages, including those with deletions. Based on the ratio of dPCR droplets with single fluorescence (corresponding to amplicons with a deletion) vs those with double fluorescence (amplicons without the deletion), the proportion of VOCs in the wastewater can be estimated (Figure 1). Further, Alpha variants can be distinguished from Beta and Gamma by comparing dPCR results of the two assays targeting the spike  $\Delta 69-70$  and ORF1a  $\Delta 3675-3677$  [1].

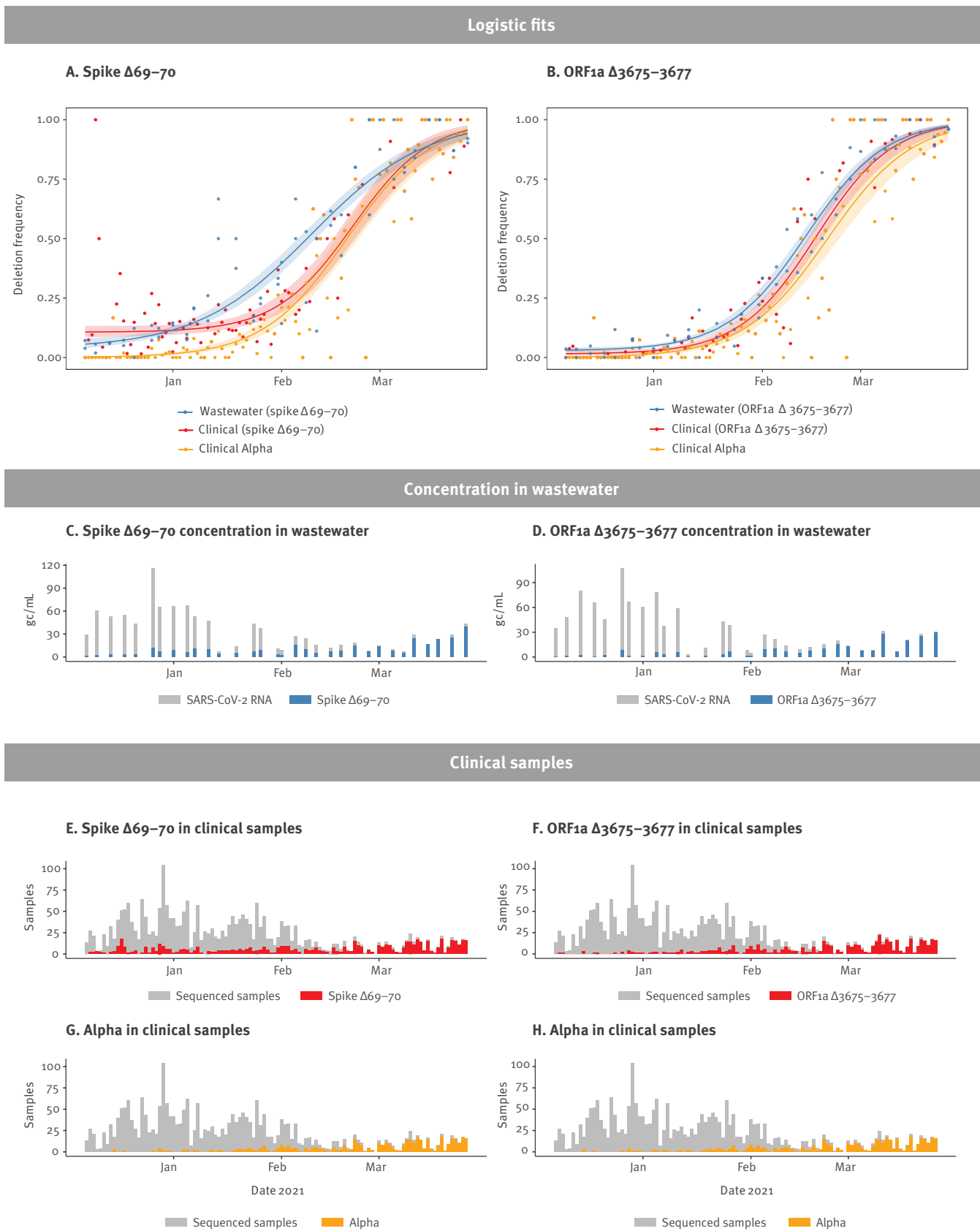
The drop-off RT-dPCR assays were performed with the Crystal Digital PCR using the Naica System (Stilla Technologies, Villejuif, France) with the qScript XLT 1-Step RT-PCR Kit (QuantaBio, Beverly, Massachusetts, United States (US)). Reactions were prepared as 27  $\mu$ L (using 5.4  $\mu$ L template and 21.6  $\mu$ L mastermix) pre-reactions, of which 25  $\mu$ L was loaded into Sapphire Chips (Stilla Technologies) for an equivalent of 5  $\mu$ L template per reaction. Mastermix composition was slightly different depending on the assay (see Supplementary Table S1 for the composition of the reactions). Templates (RNA extracts from wastewater) were diluted 1:2 to maximise the quantification of amplicons per reaction. Although dilution reduces the concentration of template in the sample, it also dilutes out inhibitory substances (the effect of different dilutions on the quantified number of amplicons per reaction is exemplified in Supplementary Figure S1). Any measurement was performed in duplicate. Thermocycler conditions included partitioning droplets (40 °C for 12 min), reverse transcription (55 °C for 30 min) and polymerase activation (95 °C for 1 min), followed by 40 cycles of denaturation (95 °C for 10 s) and annealing/extension (55 °C for 30 s).

### Benchmarking the drop-off RT-dPCR assay

We first tested the performance of the drop-off RT-dPCR assay against RNA extracts of the SARS-CoV-2 Wuhan-Hu-1 lineage (representing the wild-type sample) and of the Alpha variant (representing the VOC sample). Viral strains were obtained from the Spiez Laboratory as reported in Bechtold et al. [15]. RNA extracts from both samples were diluted in molecular grade water to a target concentration of 200 genome copies (gc)/ $\mu$ L. Each hydrolysis probe was first tested in singleplex assays. Then the universal and the deletion probe for each amplicon were combined to yield two duplex assays targeting amplicons inclusive of either spike  $\Delta 69-70$  or ORF1a  $\Delta 3675-3677$ . To assess matrix effects on assay performance, we tested both duplex assays using samples of wild type or Alpha RNA diluted in RNA extract from wastewater previously tested to contain no SARS-CoV-2 RNA. We did not use RNA from any SARS-CoV-2 Beta or Gamma strains since the method is based on a previous paper by Vogels et al., where a variety of different SARS-CoV-2 genomes have been tested, including Beta and Gamma [1].

**FIGURE 2**

Proportion of SARS-CoV-2 deletion mutants in wastewater and clinical samples and proportion of Alpha lineage in clinical samples, Zurich, Switzerland, December 2020–March 2021 (n = 32)



SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Panels A and B: Fitted curves: three-parametric logistic fits for the deletions in wastewater and clinical data and two-parametric logistic fit for the Alpha clinical data; shaded areas: 95% confidence bands. Panels C and D: concentration of SARS-CoV-2. Panels E–H: Number of Zurich sequenced clinical samples.

The performance of the duplex assays to estimate the ratio of VOCs in a mixed sample was tested in samples containing both wild-type and Alpha RNA with the following proportions of variant: 0.01, 0.02, 0.10, 0.50, 0.90, 0.98 and 0.99, at a total target concentration of 100 gc/μL. Dilutions were performed in molecular grade water.

The limit of quantification (LOQ) was determined by measuring 10 replicates of 25, 30 and 40 gc/reaction using the wild-type RNA. The LOQ was defined as the lowest value where the relative standard deviation is smaller than 25% [16]. For spike Δ69–70, the LOQ was at 40 gc/reaction and for ORF1a Δ3675–3677 at 25 gc/reaction.

### Measuring variants of concern in wastewater

We applied the drop-off RT-dPCR assays to 32 raw influent samples (24 h flow proportional composite) collected from the wastewater treatment plant Werdhölzli (serving ca 471,000 people in Zurich, Switzerland) between 7 December 2020 and 26 March 2021. Samples were stored at 4°C for on average 2–3 days but some as long as 8 days as they were shipped twice per week from the treatment plant to Eawag (Duebendorf, Switzerland), where they were immediately concentrated and the RNA was extracted. RNA extracts were stored at –80°C and processed only after assay development (which was 4 months after the first wastewater samples were concentrated and extracted). Throughout this period, a subset of 50 mL wastewater samples were spiked with murine hepatitis virus at concentrations of approximately 10<sup>6</sup> gc/50 mL to estimate virus recovery as previously described [17].

Sample concentration and RNA extraction followed a previously reported protocol [17,18]. In brief, raw influent (50 mL) was clarified by centrifugation or filtration (0.22 μM filter), then samples were concentrated using ultrafiltration (10 kDa Centricon Plus-70, Millipore, Burlington, Massachusetts, US) resulting in volumes between 150 μL and 300 μL. We then extracted RNA from the full volume of the concentrate using the QiaAmp Viral RNA MiniKit (Qiagen, Hilden, Germany), eluted it in 80 μL and, for samples after 14 January, further purified it using OneStep PCR Inhibitor Removal Kits (Zymo Research, Irvine, California, US) to reduce assay inhibition. Within the scope of the measurements of SARS-CoV-2 N1 gene marker, in which SARS-CoV-2 load is estimated, we measured assay inhibition using synthetic SARS-CoV-2 RNA as a spiked internal control [18]. However, as previously discussed, samples in this study were instead diluted 1:2 to optimise the quantification of target copies per reaction, not to accurately estimate total SARS-CoV-2 concentrations in the sample (see Supplementary Figure S1 for number of target copies quantified in the sample at various dilutions). The dMIQE checklist is available in Supplementary Table S2 [19].

### Data analysis

The statistical methods developed and used for analysing the present data are made available in an R package (R Foundation, Vienna, Austria) ‘WWdPCR’ (<https://github.com/cbg-ethz/WWdPCR>). The proportion of amplicons with deletions was calculated for each assay by accounting for the distribution of droplets with no, single and double fluorescence (see Supplement section S1 and Figure S3 for a description of how fluorescence counts were extracted). Statistical methodology to model droplet counts in singleplex dPCR is well developed [20–22]. However, to the best of our knowledge, no proper likelihood model exists for counts generated by duplex drop-off assays. In assays with multiple targets such as duplex assays, positive counts for each probe are typically assumed to be independent [20,23,24]. In a drop-off assay, this assumption is flawed as the number of viral particles positive for the drop-off probe trapped in a partition cannot exceed the number of viral particles positive for the conserved probe. When the relative proportion of variant RNA in the sample is very low (which is common in surveillance of wastewater for SARS-CoV-2) or very high, confidence intervals based on this assumption will be vastly flawed. Here, we developed a likelihood model for the duplex drop-off assay by dropping this problematic assumption. As for singleplex assays [20–22], we assumed that the number  $M_i$  of virus RNA genome copies in a droplet  $i$  is Poisson-distributed with mean  $\lambda = Cv$ , where  $C$  is the concentration of SARS-CoV-2 RNA in genome copies per volume and  $v$  is the droplet volume [20–22]. We further assumed that the number  $N_i$  of virus particles without the deletion (double-positive droplets) in a droplet  $i$  is binomially distributed with sample size  $M_i$  such that  $r_{wt}$  corresponds to the relative frequency of the wild type in the viral population.

Assuming independent and identically distributed fluorescence signals of droplets, the counts  $X_0$  of double-negative droplets (no virus RNA),  $X_1$  of droplets positive for the universal probe but not for the deletion probe (variant RNA), and  $X_2$  of double-positive droplets (wild type RNA) follow a multinomial distribution with probabilities of success:

$$p_0 = e^{-\lambda},$$

$$p_1 = e^{-r_{wt}\lambda} - e^{-\lambda},$$

$$p_2 = 1 - e^{-r_{wt}\lambda}$$

and sample size:

$$n_{tot} = x_0 + x_1 + x_2$$

corresponding to the total number of droplets (see Supplement section S2 for details of derivation).

**TABLE 2**

Logistic growth model parameter estimates for the prevalence of SARS-CoV-2 spike Δ69–70 and ORF1a Δ3675–3677 in wastewater (n = 32), Swiss clinical (n = 8,877) and Zurich clinical samples (n = 2,497), and of Alpha variants in Swiss and Zurich clinical data, Switzerland, 7 December 2020–26 March 2021

Target	Sample type	Growth rate (a)	Time to maximum growth (t <sub>o</sub> )	Background prevalence (c)	Transmission fitness advantage (f <sub>d</sub> )
Alpha	Switzerland (clinical)	0.07 (0.07–0.08)	66.1 (64.4–67.8) 11 Feb (9–13)	0 (fixed)	0.42 (0.38–0.46)
	Zurich (clinical)	0.08 (0.07–0.1)	74.6 (70.0–79.2) 20 Feb (15–24)	0 (fixed)	0.49 (0.38–0.61)
Spike Δ69–70	Wastewater	0.06 (0.06–0.07)	64.5 (62.0–67.1) 10 Feb (7–12)	0.04 (0.01–0.06)	0.34 (0.30–0.39)
	Switzerland (clinical)	0.07 (0.06–0.08)	65.5 (63.5–67.6) 11 Feb (8–12)	0.07 (0.04–0.09)	0.41 (0.36–0.46)
	Zurich (clinical)	0.09 (0.07–0.11)	76.3 (71.8–80.9) 21 Feb (17–25)	0.11 (0.07–0.14)	0.55 (0.39–0.73)
ORF1a Δ3675–3677	Wastewater	0.09 (0.08–0.09)	68.5 (67.4–69.6) 14 Feb (12–14)	0.03 (0.02–0.04)	0.53 (0.49–0.57)
	Switzerland (clinical)	0.09 (0.08–0.11)	68.7 (66.1–71.2) 14 Feb (11–13)	0.03 (0–0.05)	0.56 (0.45–0.67)
	Zurich (clinical)	0.09 (0.08–0.1)	71.3 (68.8–73.7) 17 Feb (15–19)	0.01 (0–0.03)	0.55 (0.46–0.65)

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Logistic model parameter estimates for the prevalence of spike Δ69–70 and ORF1a Δ3675–3677 in wastewater data, Swiss clinical data and Zurich clinical data are from a three-parametric logistic model (3PL) to account for possible background prevalence of the mutations.

Logistic model parameter estimates for the prevalence of Alpha variants in Swiss and Zurich clinical data are from a two-parametric logistic model (2PL). Values are maximum likelihood estimates and Wald 95% confidence intervals of the growth rate *a*, midpoint *t<sub>o</sub>* (in days after 7 December 2020 and corresponding dates) and (in the case of 3PL) background prevalence *c*. Values for the rate parameter *a* are also shown transformed (along with their confidence intervals) into an estimate of the transmission fitness advantage *f<sub>d</sub>* assuming the discrete-time growth model found in Chen et al. [12]. 3PL models are shown only when the inclusion of a third parameter (background prevalence) was statistically significant (Supplementary Table S2).

From this likelihood, we find maximum likelihood estimators:

$$\widehat{r}_{wt} = \frac{\log\left(\frac{x_0+x_1}{n_{tot}}\right)}{\log\left(\frac{x_0}{n_{tot}}\right)}$$

and

$$\widehat{\lambda} = -\log\left(\frac{x_0}{n_{tot}}\right).$$

By invariance of the likelihood, the maximum likelihood estimator of the proportion of the mutant allele *r<sub>mut</sub>* in the sample is:

$$\widehat{r}_{mut} = 1 - \widehat{r}_{wt} = \frac{\log\left(\frac{x_0}{x_0+x_1}\right)}{\log\left(\frac{x_0}{n_{tot}}\right)}.$$

The temporal increase in the proportion of amplicons with deletions in wastewater, indicative of VOCs, was compared with the temporal increase of the deletions and Alpha in clinical samples in Canton Zurich (region with 1.5 million inhabitants) and Switzerland. In analysing clinical samples from Zurich and Switzerland, we assumed that the proportion of clinical sample genomes that are Alpha or bear a signature mutation were binomially distributed, such that their empirical frequency is the maximum likelihood estimator for their prevalence in the population.

Following Chen et al. [12], we assumed that the relative frequency (*t*) of a VOC follows a logistic growth with rate *a* and inflection point *t<sub>o</sub>* (see Supplement section S3 for further discussion of model assumptions), such that:

$$r_{VOC}(t) = \frac{e^{a(t-t_0)}}{1 + e^{a(t-t_0)}}$$

Notably, here and throughout,  $t_0$  is defined as days since 7 December 2020, the first day we included in analysis.

As some lineages with no reproductive fitness advantage already present in Zurich can share mutations with the VOC under investigation, we added a background constant prevalence parameter  $c$  when modelling the relative frequency  $r_{mut}(t)$  of a mutation, such that it follows [25]:

$$r_{mut}(t) = c + (1 - c) \frac{e^{a(t-t_0)}}{1 + e^{a(t-t_0)}}$$

Borrowing from the item response theory literature of psychometrics, we will refer to these models as the two-parameter logistic (2PL) and the three-parametric logistic (3PL) models [25]. We assumed that each observation of the proportion of a mutation is an independent observation. Therefore, a best fit can be found by maximising the joint (log-)likelihood of either the binomial counts in clinical data or the multinomial counts in wastewater data. The L-BFGS-B algorithm implemented in the R v4.0.5 package ‘stats’ was used for the optimisation and estimation of the Fisher information matrix. We then computed Wald confidence intervals for the parameter estimates using the standard errors from quasi-binomial and quasi-multinomial distributions to account for over-/underdispersion [26,27]. The estimates of the growth rate parameter  $a$  were transformed along with their confidence intervals into transmission fitness advantage  $f_d$ , assuming the discrete-time logistic growth model of Chen et al. and a generation time  $g = 4.8$  days, such that  $f_d = e^{ag} - 1$  [12]. Wald confidence bands for the fitted values were calculated using the Delta method [27] for the logit-transformed fitted values and then back-transformed to the original scale, so as to ensure that they were restricted to the interval [0,1]. Approximate confidence bands were also calculated on the logit scale and then back-transformed. For each time series, both 3PL and 2PL were fit to the data and then compared by way of a likelihood ratio test based on quasi-likelihood [26], and the 3PL model was retained only if  $p < 0.05$ . The assumption of constant overdispersion of the quasi-likelihood models was checked graphically (Supplementary Figure S5).

### Ethical statement

Ethical approval was not needed for this study. SARS-CoV-2 RNA detected from wastewater in a large catchment area (471,000 people) provides anonymised data. Clinical data used for comparison were obtained from the publicly available database accessible through CovSpectrum (cov-spectrum.org) [28].

## Results

### Benchmarking the drop-off RT-dPCR assay

Both drop-off assays targeting the spike  $\Delta 69-70$  and the ORF1a  $\Delta 3675-3677$  performed as expected in singleplex and duplex assays. Specifically, the universal probe was detected in singleplex and duplex assays when using RNA from both the Alpha VOC sample and the wild-type strain. Similarly, the probe targeting only the wild-type strain (deletion probe) was only detected in the wild-type strain. Notably, when the assays were applied to the wild-type RNA, the universal probe and the deletion probe reported nearly identical concentrations. In the spike  $\Delta 69-70$  assay, the concentrations measured were 162.3 gc/ $\mu$ L for the universal probe and 162.8 gc/ $\mu$ L for the deletion probe. For the ORF1a  $\Delta 3675-3677$  assay, the concentration measured was 155.2 gc/ $\mu$ L for both probes. When testing the ratio of the Alpha RNA to the wild-type strain, both duplex assays performed as expected. The expected ratio and the measured ratio were linearly correlated (see Supplementary Figure S6). The exception was one replicate of the sample in the spike  $\Delta 69-70$  assay with a variant-to-wild type RNA ratio of 1:50, in which the variant RNA was not detected.

The assay performance was the same in wastewater matrix, with both the universal and the deletion probes detectable in the wild-type strain RNA and only the universal probe detectable in the VOC Alpha RNA. Concentrations were nearly identical for the wild-type RNA for the universal (43.3 gc/ $\mu$ L) and deletion (42.9 gc/ $\mu$ L) probes in the spike  $\Delta 69-70$  assay and for the universal (42.2 gc/ $\mu$ L) and deletion (42.2 gc/ $\mu$ L) probes in the ORF1a  $\Delta 3675-3677$  assay.

### Measuring variants of concern in Zurich wastewater

In the wastewater samples, concentrations reported by the universal probes (per mL wastewater) ranged from 2.0 gc/mL to 60.6 gc/mL for the spike  $\Delta 69-70$  assay and from 0.50 gc/mL to 65.0 gc/mL for the ORF1a  $\Delta 3675-3677$  assay (Figure 2). Concentrations of amplicons with deletions, estimated based on the number of single-positive droplets, increased over time and ranged from 0.6 gc/mL to 20.4 gc/mL for the spike  $\Delta 69-70$  assay and from undetectable to 15.8 gc/mL for the ORF1a  $\Delta 3675-3677$  assay. Notably, spike  $\Delta 69-70$  was detectable throughout the analysis period, probably due to carriage of the deletion in non-Alpha clinical samples. Recovery of MHV was 0.09% (standard deviation: 0.06%), in line with previous results [17].

Quantification of the proportion of deletions in wastewater aligned with the proportion of deletions detected in clinical samples from both Zurich (Figure 2, see Supplementary Figure S7 for the same fits with prediction bands) and Switzerland (see Supplementary Figure S4 for the comparison of wastewater samples with clinical samples throughout Switzerland and see Figure S8 for fits with prediction bands). Notably, both

deletions were detected early in wastewater, probably reflecting the presence of the deletions in variants other than Alpha, Beta and Gamma. Indeed, 3.4% of 8,709 clinical samples contained spike  $\Delta 69-70$  and 0.06% contained ORF1a  $\Delta 3675-3677$  in Switzerland before 24 December 2020 [6].

### Estimating Alpha growth rate and transmission fitness in wastewater

Accounting for these background prevalence rates in the logistic growth model, estimates for the growth rate  $a$ , time to maximum growth  $t_0$ , and transmission fitness advantage  $f_d$  were generally consistent with estimates obtained from clinical samples across Switzerland (Table 2). Specifically, the growth rate estimate from wastewater was 0.06 (95% confidence interval (CI): 0.06–0.07) for spike  $\Delta 69-70$  compared with an estimate of 0.09 (95% CI: 0.07–0.11) from Zurich clinical samples. Similarly, the growth rate estimate was 0.09 (95% CI: 0.08–0.09) for ORF1a  $\Delta 3675-3677$  in wastewater compared with 0.09 (95% CI: 0.08–0.10) from Zurich clinical samples. Converting growth rates to transmission fitness advantage further highlighted the similarities, with wastewater estimates for spike  $\Delta 69-70$  of 0.34 (95% CI: 0.30–0.39) and for ORF1a  $\Delta 3675-3677$  of 0.53 (95% CI: 0.39–0.73). These estimates of transmission fitness advantage are similar to those based on clinical samples, using the fractions of both deletions and using the fraction of Alpha variant cases. Table 2 shows that confidence intervals for the parameters of epidemiological models fitted on both types of data are for the greater part overlapping.

Estimates of time to maximum growth estimates were also generally similar when based on clinical or wastewater samples, falling within a range of less than 2 weeks in mid-February (10–21 February 2021). Here, estimates from Zurich clinical samples lagged behind both wastewater and Swiss clinical samples for both assays, with greater deviation for the spike  $\Delta 69-70$  deletion (Table 2). Similarly, the estimated background prevalences  $c$  of the mutations were in general consistent between models based on the wastewater data and on the clinical data. The inclusion of a background prevalence parameter statistically significantly improved the fit in all models based on prevalence of deletions, but not in any model based on the prevalence of the Alpha variant in clinical data (Supplementary Table S2 provides additional detail on model parameter estimates for logistic growth models based on wastewater and clinical samples using both two parametric and three parametric fits).

### Discussion

Drop-off RT-dPCR assays for targeted detection of signature mutations in wastewater with a catchment of ca 471,000 inhabitants show temporal trends consistent with sequencing data from clinical samples from the same canton (1.5 million inhabitants). By incorporating non-competitive probes on a single amplicon, proportions of the mutation in the virus population

can be directly estimated with a single assay that is not biased by inter-assay variation in quantification. The analysis of wastewater in this study included 32 samples, which contrasts with 2,497 clinical samples collected in Canton Zurich and sequenced during the period of study (7 December 2020 to 26 March 2021). The demonstrated approach provides a method of tracking signature mutations in a community at lower cost and with faster turnaround than clinical or wastewater sequencing, and can be adapted to screen for other signature mutations of other VOCs.

Temporal trends in signature mutations (here, ORF1a  $\Delta 3675-3677$  and spike  $\Delta 69-70$ ) of VOCs tracked in wastewater using dPCR can rapidly inform variant transmissibility. Point estimates for advantages in growth rate and transmission fitness derived from wastewater were generally consistent with point estimates derived from clinical samples: the 95% confidence intervals overlapped as shown in the confidence intervals of the parameter estimates. Because dPCR sample analysis is faster than whole-genome sequencing, the workflow presented here allows insights into transmissibility of novel variants once signature mutations are identified.

Notably, we found that the proportion of variants in wastewater increased earlier than in clinical samples in Zurich. In wastewater samples, the maximum growth rate of ORF1a  $\Delta 3675-3677$  occurred 2.8 days earlier than in clinical samples, and the maximum growth rate of spike  $\Delta 69-70$  occurred 11.8 days earlier than in clinical samples. There are a number of possible explanations for the discrepancy in dates between clinical samples and wastewater samples. Firstly, the dates of the two types of samples represent distinct events, with the wastewater samples representing the 24 h period when the wastewater sample was collected and the dates of the clinical samples representing the date a person visits a testing facility. The dynamics of clinical testing may not align with the dynamics of shedding into wastewater. Previous work comparing temporal trends in cases with temporal trends in Zurich wastewater suggests this may only partially contribute to the discrepancy. Wastewater signals for SARS-CoV-2 were observed to be at most 1 day earlier than clinical case data, based on the average delay between infection and date of testing in Zurich (8.1 days) compared with the estimated delay between infection and shedding of SARS-CoV-2 RNA in wastewater (7–11 days) [18]. A second explanation may be that shedding (i.e. concentration, duration) of SARS-CoV-2 RNA into the wastewater may vary by variant. In clinical samples of SARS-CoV-2 RNA measurements from nose/throat samples, people infected with the Alpha variant shed substantially higher viral concentrations than people infected with non-Alpha variants [29]. If more Alpha SARS-CoV-2 RNA is shed into wastewater per person infected as compared with non-Alpha variants, then the proportion of variant amplicons in wastewater would be skewed towards Alpha and would not be an accurate estimate of the proportion of infections with the Alpha variant.



Although there is a variant-specific impact on shedding observed from nose/throat samples, it is largely unknown if there is also a variant-specific impact on faecal shedding. Variants may influence not only the magnitude of enteric shedding, but also the proportion of infected people with enteric shedding. Theoretically, wastewater-based estimates on the frequency of VOCs in the population could be corrected to account for variant-specific differences in enteric shedding, but this would require reliable estimates of enteric shedding load profiles for each variant, which are currently lacking.

We applied the drop-off RT-dPCR assay here for signature mutations of SARS-CoV-2 Alpha, Gamma and Beta, but the assay is readily adaptable to screen for other signature mutations beyond those presented here. Targeting signature deletions allows for the design of specific probes annealing to the deleted region to distinguish target variants from other lineages. However, probes providing specificity to other mutations, including single nucleotide polymorphisms, could be targeted by future drop-off assays. Examples include mutations that may influence not only transmissibility but also virulence and/or vaccine-induced antibody escape [13].

The methodology has some limitations. Notably, wastewater-based estimation of the proportion of signature mutations in communities requires sufficient concentrations of SARS-CoV-2 RNA such that relative proportions of target mutations can be quantified. For example, lags in sample processing may have decreased concentrations of SARS-CoV-2 RNA in the samples [17]. Methodological improvements to increase the quantity of SARS-CoV-2 RNA extracted from wastewater are needed to apply this method in scenarios of low viral RNA concentrations in wastewater (corresponding to few clinical cases in the catchment) [30]. Although concentrations of the deletion amplicons in the wastewater often are below the limit of quantification (defined as the lowest concentration with a coefficient of variation  $\leq 25\%$ ), transmission fitness estimates remain aligned with estimates from clinical samples. This finding suggests that low SARS-CoV-2 RNA concentrations in wastewater may nevertheless be sufficient to infer epidemiological data when analysed as a time series. In addition, the methodology assumes that the universal probe targets a region of the amplicon conserved across all strains and therefore is an accurate indicator of the total number of SARS-CoV-2 amplicons in the sample. Strain-specific variation in the efficiency of amplification of the universal probe, for example because of mutations or differential sample inhibition [31], would result in inaccurate estimation of the total number of amplicons and corresponding inaccuracy of the proportion of the target variant. To address this, continual monitoring of clinical sample sequences for mutations in the target regions of the primers and both universal and wild-type probes is recommended.

## Conclusion

Drop-off RT-dPCR assays provide an opportunity for rapid detection and quantitative estimates of VOCs in communities, offering epidemiological insights into their introduction and transmission. Such assays offer complementary tools to sequencing analyses of VOCs which may be useful for large-scale programmes in wastewater-based epidemiology of SARS-CoV-2 such as the US National Sewer Surveillance System and the European Union Sewer Sentinel Surveillance Network.

## Acknowledgements

We thank the operators of the Zurich WWTPs for providing samples. Funding was provided through the Swiss National Science foundation (Special Call on Coronaviruses (31CA30\_196538), Federal Office of Public Health, Eawag and EPFL discretionary funding. XFC was a fellow of the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 754462. An earlier version of this article appears as a preprint under MedRxiv under <https://doi.org/10.1101/2021.08.22.21262024>.

## Conflict of interest

None declared.

## Authors' contributions

LC, TK, CO, and TRJ, conceived the study; DD and NB developed and applied statistical and analytical methods; LC, TS, AJD, PG, AK, ES, XFC, TK, CO, and TRJ developed experimental protocols and performed wastewater sampling; XFC, NB, TK, CO, and TRJ supervised the study and secured funding; LC, DD, and TRJ wrote the original draft; all authors reviewed and approved the final manuscript.

## References

1. Vogels CBF, Breban MI, Ott IM, Alpert T, Petrone ME, Watkins AE, et al. Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. *PLoS Biol.* 2021;19(5):e3001236. <https://doi.org/10.1371/journal.pbio.3001236> PMID: 33961632
2. Bedotto M, Fournier P-E, Houhamdi L, Levasseur A, Delerce J, Pinault L, et al. Implementation of an in-house real-time reverse transcription-PCR assay for the rapid detection of the SARS-CoV-2 Marseille-4 variant. *J Clin Virol.* 2021;139:104814. <https://doi.org/10.1016/j.jcv.2021.104814> PMID: 33836314
3. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA.* 2020;323(18):1843-4. <https://doi.org/10.1001/jama.2020.3786> PMID: 32159775
4. Jones DL, Baluja MQ, Graham DW, Corbishley A, McDonald JE, Malham SK, et al. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Sci Total Environ.* 2020;749:141364. <https://doi.org/10.1016/j.scitotenv.2020.141364> PMID: 32836117
5. Crits-Christoph A, Kantor RS, Olm MR, Whitney ON, Al-Shayeb B, Lou YC, et al. Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 Variants. *MBio.* 2021;12(1):e02703-20. <https://doi.org/10.1128/mBio.02703-20> PMID: 33468686
6. Jahn K, Dreifuss D, Topolsky I, Kull A, Ganesanandamoorthy P, Fernandez-Cassi X, et al. Detection of SARS-CoV-2 variants in Switzerland by genomic analysis of wastewater samples. *medRxiv.* 2021.01.08.21249379. <https://doi.org/10.1101/2021.01.08.21249379> <https://doi.org/10.1101/2021.01.08.21249379>
7. Wilton T, Bujaki E, Klapsa D, Majumdar M, Zambon M, Fritzsche M, et al. Rapid increase of SARS-CoV-2 variant B.1.1.7 detected

- in sewage samples from England between October 2020 and January 2021. *mSystems*. 2021;6(3):e0035321. <https://doi.org/10.1128/mSystems.00353-21> PMID: 34128696
8. Chaintoutis SC, Chassalevris T, Tsiolas G, Balaska S, Vlatakis I, Mouchtaropoulou E, et al. A one-step real-time RT-PCR assay for simultaneous typing of SARS-CoV-2 mutations associated with the E484K and N501Y spike protein amino-acid substitutions. *J Virol Methods*. 2021;296:114242. <https://doi.org/10.1016/j.jviromet.2021.114242> PMID: 34274369
  9. Wurtzer S, Waldman P, Levert M, Mouchel JM, Gorgé O, Boni M, et al. Monitoring the propagation of SARS-CoV-2 variants by tracking identified mutation in wastewater using specific RT-qPCR. *medRxiv*. 2021.03.10.21253291 <https://doi.org/10.1101/2021.03.10.21253291> <https://doi.org/10.1101/2021.03.10.21253291>
  10. Lee WL, Imakaev M, Armas F, McElroy KA, Gu X, Duvallet C, et al. Quantitative SARS-CoV-2 Alpha variant B.1.1.7 tracking in wastewater by allele-specific RT-qPCR. *Environ Sci Technol Lett*. 2021;8(8):675-82. <https://doi.org/10.1021/acs.estlett.1c00375>
  11. Heijnen L, Elsinga G, de Graaf M, Molenkamp R, Koopmans MPG, Medema G. Droplet digital RT-PCR to detect SARS-CoV-2 signature mutations of variants of concern in wastewater. *Sci Total Environ*. 2021;799:149456. <https://doi.org/10.1016/j.scitotenv.2021.149456> PMID: 34371414
  12. Chen C, Nadeau SA, Topolsky I, Manceau M, Huisman JS, Jablonski KP, et al. Quantification of the spread of SARS-CoV-2 variant B.1.1.7 in Switzerland. *Epidemics*. 2021;37:100480. <https://doi.org/10.1016/j.epidem.2021.100480> PMID: 34488035
  13. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. 2021;19(7):409-24. <https://doi.org/10.1038/s41579-021-00573-0> PMID: 34075212
  14. Meng B, Kemp SA, Papa G, Datir R, Ferreira IATM, Marelli S, et al. Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep*. 2021;35(13):109292. <https://doi.org/10.1016/j.celrep.2021.109292> PMID: 34166617
  15. Bechtold P, Wagner P, Hosch S, Siegrist D, Ruiz-Serrano A, Gregorini M, et al. Rapid identification of SARS-CoV-2 variants of concern using a portable peakPCR Platform. *Anal Chem*. 2021;93(49):16350-9. <https://doi.org/10.1021/acs.analchem.1c02368> PMID: 34852455
  16. Zhu K, Suttner B, Pickering A, Konstantinidis KT, Brown J. A novel droplet digital PCR human mtDNA assay for fecal source tracking. *Water Res*. 2020;183:116085. <https://doi.org/10.1016/j.watres.2020.116085> PMID: 32750535
  17. Fernandez-Cassi X, Scheidegger A, Bänziger C, Cariti F, Tuñas Corzon A, Ganesanandamoorthy P, et al. Wastewater monitoring outperforms case numbers as a tool to track COVID-19 incidence dynamics when test positivity rates are high. *Water Res*. 2021;200:117252. <https://doi.org/10.1016/j.watres.2021.117252> PMID: 34048984
  18. Huisman JS, Scire J, Caduff L, Fernandez-Cassi X, Ganesanandamoorthy P, Kull A, et al. Wastewater-based estimation of the effective reproductive number of SARS-CoV-2. *medRxiv*. 2021.04.29.21255961. <https://doi.org/10.1101/2021.04.29.21255961> <https://doi.org/10.1101/2021.04.29.21255961>
  19. Huggett JF, dMIQE Group. The digital MIQE guidelines update: minimum information for publication of quantitative digital PCR experiments for 2020. *Clin Chem*. 2020;66(8):1012-29. <https://doi.org/10.1093/clinchem/hvaa125> PMID: 32746458
  20. Dube S, Qin J, Ramakrishnan R. Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. *PLoS One*. 2008;3(8):e2876. <https://doi.org/10.1371/journal.pone.0002876> PMID: 18682853
  21. Basu AS. Digital assays part I: partitioning statistics and digital PCR. *SLAS Technol*. 2017;22(4):369-86. <https://doi.org/10.1177/2472630317705680> PMID: 28448765
  22. Dorazio RM, Hunter ME. Statistical models for the analysis and design of digital polymerase chain reaction (dPCR) experiments. *Anal Chem*. 2015;87(21):10886-93. <https://doi.org/10.1021/acs.analchem.5b02429> PMID: 26436653
  23. Whale AS, Huggett JF, Tzonev S. Fundamentals of multiplexing with digital PCR. *Biomol Detect Quantif*. 2016;10:15-23. <https://doi.org/10.1016/j.bdq.2016.05.002> PMID: 27990345
  24. Milbury CA, Zhong Q, Lin J, Williams M, Olson J, Link DR, et al. Determining lower limits of detection of digital PCR assays for cancer-related gene mutations. *Biomol Detect Quantif*. 2014;1(1):8-22. <https://doi.org/10.1016/j.bdq.2014.08.001> PMID: 27920993
  25. Baker FB, Kim S-H. Item response theory: parameter estimation techniques, second edition. CRC Press. 2nd ed. Boca Raton: CRC Press; 2004 <https://doi.org/10.1201/9781482276725>
  26. McCullagh P, Nelder JA. Generalized linear models. 2nd ed. New York: Routledge; 2019. <https://doi.org/10.1201/9780203753736>
  27. Held L, Sabanés Bové D. Applied statistical inference: likelihood and bayes. Berlin, Heidelberg: Springer; 2014. <http://dx.doi.org/10.1007/978-3-642-37887-4>
  28. Chen C, Nadeau S, Yared M, Voinov P, Xie N, Roemer C, et al. CoV-spectrum: analysis of globally shared SARS-CoV-2 data to identify and characterize new variants. *Bioinformatics*. 2021;btab856.
  29. Jones TC, Biele G, Mühlemann B, Veith T, Schneider J, Beheim-Schwarzbach J, et al. Estimating infectiousness throughout SARS-CoV-2 infection course. *Science*. 2021;373(6551):eabi5273. <https://doi.org/10.1126/science.abi5273> PMID: 34035154
  30. Philo SE, Keim EK, Swanstrom R, Ong AQW, Burnor EA, Kossik AL, et al. A comparison of SARS-CoV-2 wastewater concentration methods for environmental surveillance. *Sci Total Environ*. 2021;760:144215. <https://doi.org/10.1016/j.scitotenv.2020.144215> PMID: 33340739
  31. Huggett JF, Novak T, Garson JA, Green C, Morris-Jones SD, Miller RF, et al. Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon. *BMC Res Notes*. 2008;1(1):70. <https://doi.org/10.1186/1756-0500-1-70> PMID: 18755023

### License, supplementary material and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2022.