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.
Monitoring COVID-19 spread in Prague local neighborhoods based on the presence of SARS-CoV-2 RNA in wastewater collected throughout the sewer network

Kamila Zdenkova a,*, Jana Bartackova b, Elisha Cermakova a, Katerina Demnerova a, Alzbeta Dostalkova c, Vaclav Janda b, Jiri Jarkovsky d, Marco Antonio Lopez Marin b, Zuzana Novakova e, Michaela Rumlova c, Jana Rihova Ambrozova b, Klara Skodakova b, Iva Swierczkova f, Petr Sykora e, Dana Vejmelkova b, Jiri Wanner b, Jan Bartacek b

a Department of Biochemistry and Microbiology, University of Chemistry and Technology Prague, Czechia
b Department of Water Technology and Environmental Engineering, University of Chemistry and Technology Prague, Czechia
c Department of Biotechnology, University of Chemistry and Technology Prague, Czechia
d Institute of Health Information and Statistics of the Czech Republic, Czechia
e Prazske vodovody a kanalizace, a.s., Czechia
f Military Health Institute, Military Medical Agency, Czechia

ARTICLE INFO

Keywords:
SARS-CoV-2
Wastewater-based epidemiology
Prague
RT-mqPCR
COVID-19 epidemic

ABSTRACT

Many reports have documented that the presence of SARS-CoV-2 RNA in the influents of municipal wastewater treatment plants (WWTP) correlates with the actual epidemic situation in a given city. However, few data have been reported thus far on measurements upstream of WWTPs, i.e. throughout the sewer network. In this study, the monitoring of the presence of SARS-CoV-2 RNA in Prague wastewater was carried out at selected locations of the Prague sewer network from August 2020 through May 2021. Various locations such as residential areas of various sizes, hospitals, city center areas, student dormitories, transportation hubs (airport, bus terminal), and commercial areas were monitored together with four of the main Prague sewers. The presence of SARS-CoV-2 RNA was determined by reverse transcription – multiplex quantitative polymerase chain reaction (RT-mqPCR) after the precipitation of nucleic acids with PEG 8,000 and RNA isolation with TRIzol™ Reagent. The number of copies of the gene encoding SARS-CoV-2 nucleocapsid (N1) per liter of wastewater was compared with the number of officially registered COVID-19 cases in Prague. Although the data obtained by sampling wastewater from the major Prague sewers were more consistent than those obtained from the small sewers, the correlation between wastewater-based and clinical-testing data was also good for the residential areas with more than 7,000 registered inhabitants. It was shown that monitoring SARS-CoV-2 RNA in wastewater sampled from small sewers could identify isolated occurrences of COVID-19-positive cases in local neighborhoods. This can be very valuable while tracking COVID-19 hotspots within large cities.

1. Introduction

Even though the first wave of the COVID-19 pandemic hit Czechia mildly from March to May 2020, the numbers of patients rose sharply later in October. Since then, the Czech Republic remained one of the most severely affected countries worldwide until April 2021. At the same time, the Czech population was relatively reluctant to undergo clinical testing, resulting in an extremely high rate of positively tested patients, 30 to 40% of which were proved to be COVID-positive. This situation led to a health crisis, which was extremely difficult to manage by the state authorities. In such a situation, an accurate and timely system for the monitoring of the pandemic is critically needed.

Since the first papers by Medema et al. (2020) reporting the possibility of detecting the presence of SARS-CoV-2 (the virus responsible for the COVID-19 disease) in sewage, hundreds of papers describing the correlation between viral RNA in wastewater and the COVID-19 epidemic have been published. However, most of these studies have focused on measurements downstream of WWTPs, i.e. in the effluents. Few data have been reported thus far on measurements upstream of WWTPs, i.e. throughout the sewer network. In this study, the monitoring of the presence of SARS-CoV-2 RNA in Prague wastewater was carried out at selected locations of the Prague sewer network from August 2020 through May 2021. Various locations such as residential areas of various sizes, hospitals, city center areas, student dormitories, transportation hubs (airport, bus terminal), and commercial areas were monitored together with four of the main Prague sewers. The presence of SARS-CoV-2 RNA was determined by reverse transcription – multiplex quantitative polymerase chain reaction (RT-mqPCR) after the precipitation of nucleic acids with PEG 8,000 and RNA isolation with TRIzol™ Reagent. The number of copies of the gene encoding SARS-CoV-2 nucleocapsid (N1) per liter of wastewater was compared with the number of officially registered COVID-19 cases in Prague. Although the data obtained by sampling wastewater from the major Prague sewers were more consistent than those obtained from the small sewers, the correlation between wastewater-based and clinical-testing data was also good for the residential areas with more than 7,000 registered inhabitants. It was shown that monitoring SARS-CoV-2 RNA in wastewater sampled from small sewers could identify isolated occurrences of COVID-19-positive cases in local neighborhoods. This can be very valuable while tracking COVID-19 hotspots within large cities.
Table 1
Sampling sites monitored in this study.

| Location no. | Type of neighborhood | No. officially registered residents | Type of samples | Sampling frequency | Separate sewer for storm water |
|--------------|----------------------|-------------------------------------|-----------------|-------------------|-------------------------------|
| 1            | Family houses 1      | 105                                 | grab            | Once per two weeks (September 2020) | 28 Yes |
| 2            | Family houses 2      | 974                                 | grab            | Twice per week (March-May 2020)    | 27 Yes |
| 3            | Family houses 3      | 10,152                              | grab            | 2020-FEBRUARY 2021               | 36 No  |
| 4            | Apartment buildings 1| 5,869                               | grab            | 2020-FEBRUARY 2021               | 26 No  |
| 5            | Apartment buildings 2| 7,075                               | grab            | 35 Yes                         | 35 Yes |
| 6            | Apartment buildings 3| 1,153                               | grab            | 36 Yes                         | 36 No  |
| 7            | Hospital & Residential area | 13,148     | grab            | 36 No                         | 38 No  |
| 8            | City center          | 1,628                               | grab            | 35 No                         | 37 Yes |
| 9            | Office buildings     | 99                                   | grab            | 35 No                         | 35 No  |
| 10           | Shopping mall        | 1,478                                | grab            | 35 No                         | 35 No  |
| 11           | University dormitories | 98                                  | grab            | 35 No                         | 35 No  |
| 12           | Industrial area      | 2,340                                | grab            |                | 35 No  |
| 13           | Airport              | n.a.*                                | grab/24h        | 12 Yes                        | 12 Yes |
| 14           | Airport              | n.a.*                                | grab/24h        | 19 Yes                        | 19 Yes |
| 15           | Family houses 4      | 1,789                                | grab            | 12 No                         | 12 No  |
| 16           | Apartment buildings 4| 6,457                                | grab            | 12 Yes                        | 12 Yes |
| 17           | Apartment buildings 5| 12,085                               | grab            | 12 No                         | 12 No  |
| 18           | Apartment buildings 6| 3,453                                | grab            | 12 No                         | 12 No  |
| 19           | Hospital area        | 0                                    | grab            | 12 No                         | 12 No  |
| 20           | City center 2        | 9,738                                | grab            | 12 No                         | 12 No  |
| 21           | University dormitories, Apartment houses | 1,874     | grab            | 12 Yes                        | 12 Yes |
| 22           | Bus terminal         | 1,318                                | grab            | 14 No                        | 14 No  |

Table 1 continued

| Sampling frequency | Separate sewer for storm water |
|--------------------|-------------------------------|
| 2-3 times per week | 31 No                         |
| Twice per week     | 21 No                         |
| Twice per week     | 22 No                         |
| Twice per week     | 22 No                         |

* n.a. – not available

Unfortunately, limited data have been published in this respect so far. Goncalves et al. (2021) had been detecting SARS-CoV-2 RNA in the hospital area over a short period of two weeks. Similarly, Acosta et al. (2021) monitored wastewater from three hospitals showing that viral burden correlated with increasing hospitalized cases as well as hospital-associated transmissions. Ahmed et al. (2021a) sampled four main sewers in Bangladesh, but again, few samples were analyzed (6 sampling days within one month) to assess their epidemiological value. Sagutti et al. (2021) reported results from 4 sampling sites throughout the city of Gothenburg (Sweden), but the sampling in these subareas only covered 4 weeks, and the correlation with epidemic data was unclear. Also, Baldwin et al. (2021) reported data for 4 sampling sites (pumping stations) throughout Padova (Italy) that each represented a large portion of the city. Finally, Yeager et al. (2021) proposed methodology for population-level sampling for SARS-CoV-2 RNA presence in wastewater. They reported data from 8 weeks monitoring (September to October 2020) of 17 sub-sewersheds in Jefferson County with isolated populations between 5,000 and 100,000 residents.

Number of authors have also targeted specific objects such as university dormitories and showed good correlation with the number of COVID-19 cases onsite or even predictive value of the wastewater-based data (Scott et al. 2021). Additionally, Vo et al. (2022) demonstrated successful analysis of viral variants in samples collected from dormitories.

Sofar the most detailed study using decentralized wastewater monitoring published very recently Mota et al. (2021) who monitored 17 locations within the sewer network of Belo Horizonte (Brazil) from May through August 2020 and showed that epidemic hotspots could be identified in the city based on data generated by decentralized sewage monitoring, rather than based on clinical data. The locations monitored...
in the latter study involved relatively large catchment areas representing from approximately 12 thousand to 1.4 million inhabitants with the majority of the areas representing several tens of thousands of inhabitants. More long-term data of similar nature, preferably from even smaller locations, are very much needed to complement conventional clinical surveillance of the epidemic.

From July 2020 through March 2021, we monitored the presence of SARS-CoV-2 RNA in sewage sampled biweekly at 14 locations throughout the Prague sewer network. Within these locations were residential areas with apartment buildings and family houses, commercial area, areas dominated by student dormitories, or Prague’s main airport. From late March through the end of May 2021, we expanded the monitoring to 26 locations including four of Prague’s main sewers and taking samples up to three times a week. In this paper, we report the presence of SARS-CoV-2 RNA in sewage and its correlation with the epidemic situation in Prague.

Following the preliminary study performed in summer 2020 by Mlejnkova et al. (2020) at WWTPs throughout the Czech Republic, this is the first consistent study on the presence of SARS-CoV-2 RNA in Prague wastewater. After the very recent paper by Mota et al. (2021), this is the first European study thoroughly reporting long-term data for local neighborhoods of different types within one city.

2. Materials and methods

2.1. Sampling campaigns

The sampling took place in two main periods: (1) a long-term monitoring, where grab samples were taken every other week at 14 locations (Locations 1 to 14, Table 1) within the Prague sewer network and (2) an intensive monitoring (2 to 3 samples per week) in spring 2021, which was done at the same locations as the long-term monitoring. In addition, the intensive monitoring included other eight local neighborhoods and four of Prague main sewers. At these trunk sewers, 24-hour composite samples have been collected.

2.2. Description of the neighborhoods monitored in this study

The local neighborhoods (Locations 1 to 22) monitored in this study (Fig. 1) were selected based on their epidemiological importance (Table 1). Specifically, we selected residential areas (Locations 1 to 7 and 15 to 18) of various types (family houses, apartment buildings) and size (105 to 13,148 registered inhabitants). Next to that, areas representing typical activities taking place in large cities were selected: Locations tourism and other commercial activities (Locations 8 and 20), activities bound to office areas (Location 9), shopping areas (Location 10), accommodation of university students (Location 11 and 21), and industrial areas (Location 12), and transportation hubs, i.e. Prague’s main airport (Locations 13, 14) and the main bus terminal (Location 22). Note that some of these areas inevitably combine more types of activities, e.g. Location 7 combines the highest number of inhabitants of all locations and a large hospital or Location 10 combines wastewater from shopping mall and an adjacent residential area with 1478 registered inhabitants.

It was the intention of the authors to prioritize areas with separate sewers for sewage and rainwater to minimize the dilution of sewage. However, it was not possible to observe this principle for all locations as combined sewers collect wastewater from most of inner Prague’s area. The trunk sewers (Fig. 1) were selected to cover a large portion of Prague’s wastewater (to obtain data representative for the whole of Prague) and to observe deviations between data obtained from large wastewater sources. Note that wastewater from Sewer ACK, the main influent to Prague’s WWTP, mainly represents the combination of sewers C and K, and therefore to some extent should give similar results. Besides sampling wastewater, also wastewater flow was measured at the sampling points S1–S2.
The procedure reported by Wu et al. (2020) was used. The sample was not refrigerated during transport to laboratory. First pasteurized at 60°C, composite samples (40 mL every 60 min) were taken. The samples were 10 a.m.). Bottles and subsequently delivered to the lab within one to two hours without any cooling. The sampling time was approximately constant for each location and chosen to represent the morning peak (between 7 and 10 a.m.).

From the main sewers (Locations S1–S4), 1L of 24 h time-controlled composite samples (40 mL every 60 min) were taken. The samples were not refrigerated during transport to laboratory.

2.3. Wastewater samples

At local neighborhoods (Locations 1–22), 500–1,000 mL grab samples were taken directly from the sewer into 1L sterilized polyethylene bottles and subsequently delivered to the lab within one to two hours without any cooling. The sampling time was approximately constant for each location and chosen to represent the morning peak (between 7 and 10 a.m.).

From the main sewers (Locations S1–S4), 1L of 24 h time-controlled composite samples (40 mL every 60 min) were taken. The samples were not refrigerated during transport to laboratory.

2.4. Samples processing

At the beginning of the monitoring (July through September 2020), the procedure reported by Wu et al. (2020) was used. The sample was first pasteurized at 60°C for 90 min and subsequently filtered through a 0.45 µm nitrocellulose membrane filter (Sigma Aldrich, UK). For each sample, several filters were used due to filter clogging until 40 mL of filtrate was produced. The filtrate was mixed with 4 g of polyethylene glycol (PEG) 8,000 (Sigma Aldrich, UK) and 0.9 g NaCl (Penta, Czechia) and centrifuged at 14,000 g and 4°C for 90 min and subsequently filtered through a 0.2 µm pipette, and 500 µL of TRIzol™ Reagent (Invitrogen ThermoFisher, USA) was added to the pellet. The pellet with TRIzol™ Reagent was either directly processed or kept at -80°C until RNA extraction (for max. 4 weeks period). After the optimization of the procedure (September 2020, data not shown), pasteurization was omitted and the vacuum filtration was replaced with centrifugation at 4,600 g and 4°C for 30 min. The supernatant was resuspended with 8 g of PEG 8,000 and 1.8 g NaCl and processed like in the procedure above.

2.5. RNA extraction

After PEG precipitation, TRIzol-chloroform extraction was used for total RNA isolation according to the manufacturer’s protocol (Invitrogen, ThermoFisher Scientific). Since the RNA was finally isolated from a total volume of 80 mL of wastewater, two pellets each initially obtained from 40 mL of wastewater and re-suspended in 500 µL of TRIzol™ Reagent were combined to obtain 1 mL of TRIzol™ suspension. To control the total RNA isolation process, 800 copies of spike ssRNA (EURO-019) were added directly into the sample (suspension of 1 mL of TRIzol™ Reagent and pellets from 80 mL of the sample). Then, 200 µL of chloroform (Penta, Czechia) was added, the sample was vortexed thoroughly for 15 s, and incubated for 15 min at room temperature. After incubation, the mixture was centrifuged at 12,000 g for 15 min at 4°C in a pre-cooled centrifuge. The aqueous phase containing RNA was transferred into a fresh tube and 500 µL of 2-propanol was added. After 8 min of incubation at room temperature, RNA was pelleted by centrifugation (10 min, 12,000 g, 4°C). Next, the supernatant was removed and the pellet was washed by adding 1.4 mL of 75% ethanol (Penta, Czechia). The sample was then centrifuged at 7,500 g for 5 min at 4°C. Finally, the obtained pellet of RNA was briefly dried and resuspended in 10–50 µL of nuclelease-free water (NFW, Top-Bio, Czechia).

2.6. Reverse transcription–multiplex quantitative real-time PCR (RT-mqPCR)

Reverse transcription and PCR amplification were conducted in a single tube at a reaction volume of 20 µL containing a 1x EliZyme OneS Probe Kit (Elisabeth Pharmacon; Czechia), primers (0.4 µM each) and probes (0.2 µM each), nuclelease-free water, and 5 µL of RNA in each reaction (undiluted and two-fold dilution). The sequences of all primers and probes used in the multiplex RT-qPCR are listed in Table 2. The oligonucleotides were synthesized by Metabion International AG (Pla-negg, Germany). Primers and probes used for RT-mqPCR were complementary to the SARS-CoV-2 genome. The sequences for the detection of the N1 nucleocapsid-encoding gene were adopted from the CDC protocol (CDC, 2020), for spike protein from the EURM-019 product list (JRC, 2020), and for the RNA-dependent RNA polymerase (RdRp) gene from Cormann et al. (2020). All RT-mqPCRs were performed using a 7500 Real-Time PCR system (Applied Biosystems, Foster City, USA), and the data were analyzed using the 7500 Software v2.0.6. The fluorescence channels were evaluated separately: the FAM fluorophore detected nucleocapsid and spike protein-encoding gene fragments (72 bp and 83 bp, respectively), TAMRA was used for the RdRp of SARS-CoV-2 (100 bp), and HEX was used for the RdRp of SARS-CoV-2, SARS-CoV, and bat-SARS-related coronaviruses (CoVs). Only the FAM channel was used for gene quantification. The serial dilution method was applied for inhibition testing in each sample. Undiluted RNA and 2x diluted RNA (dilution in NFW from Promega, Madison, WI, USA) were used. For the first two months, all samples were analyzed in duplicates. Later, only every tenth sample was done in duplicate. The positive control (target RNA) and no template control (NFW) were run on every RT-mqPCR plate to exclude false (negative or positive) results.

Reverse transcription and amplification were conducted in a single tube at a reaction volume of 20 µL containing a 1x EliZyme OneS Probe Kit (Elisabeth Pharmacon; Czechia), primers (0.4 µM each) and probes (0.2 µM each), nuclelease-free water, and 5 µL of RNA in each reaction (undiluted and two-fold dilution). The sequences of all primers and probes used in the multiplex RT-qPCR are listed in Table 2. The oligonucleotides were synthesized by Metabion International AG (Planegg, Germany). Primers and probes used for RT-mqPCR were complementary to the SARS-CoV-2 genome. The sequences for the detection of the N1 nucleocapsid-encoding gene were adopted from the CDC protocol (CDC, 2020), for spike protein from the EURM-019 product list (JRC, 2020), and for the RNA-dependent RNA polymerase (RdRp) gene from Cormann et al. (2020). All RT-mqPCRs were performed using a 7500 Real-Time PCR system (Applied Biosystems, Foster City, USA), and the data were analyzed using the 7500 Software v2.0.6. The fluorescence channels were evaluated separately: the FAM fluorophore detected nucleocapsid and spike protein-encoding gene fragments (72 bp and 83 bp, respectively), TAMRA was used for the RdRp of SARS-CoV-2 (100 bp), and HEX was used for the RdRp of SARS-CoV-2, SARS-CoV, and bat-SARS-related coronaviruses (CoVs). Only the FAM channel was used for gene quantification. The serial dilution method was applied for inhibition testing in each sample. Undiluted RNA and 2x diluted RNA (dilution in NFW from Promega, Madison, WI, USA) were used. For the first two months, all samples were analyzed in duplicates. Later, only every tenth sample was done in duplicate. The positive control (target RNA) and no template control (NFW) were run on every RT-mqPCR plate to exclude false (negative or positive) results.

The conditions for transcription and amplifications were as follows: reverse transcription for 10 min at 53°C, initial denaturation at 95°C for 2 min followed by 45 cycles of 5 s at 95°C and 30 s at 60°C.

The EURM-019 synthetic single-stranded RNA (synthetic ssRNA) fragments of SARS-CoV-2 were used as an in vitro-transcribed RNA standard (JRC, 2020). This universal synthetic ssRNA of 880 nts contains the target regions that can be amplified by all the RT-qPCR assays listed in Table 1. The procedures for verifying the methodology (quantification, amplification efficiency E, R² coefficient, repeatability) met the criteria of the JRC Technical report on the verification of analytical methods for GMO testing (Hougs et al. 2017). The determination of the limit of quantification (LOQ) and limit of detection (LOD) was performed by contaminating the samples with a known amount of synthetic ssRNA. The LOQ of the RT-mqPCR is 50 copies of N1 gene; calibration curves always included this standard and were constructed in all

### Table 2

Primer and probes used in this study.

| Target          | Name               | Sequence [5’-3’]                  | Size [bp] | Ref.          |
|-----------------|--------------------|-----------------------------------|-----------|---------------|
| N1 nucleocapsid | CDC_Wu_N1-F        | GCCCCAAAATACGGCAAAAT              | 72        | (CDC 2020)    |
|                 | CDC_Wu_N1-R        | TCTGGTACTGGCCAGTGGTAATC          |           |               |
|                 | CDC_Wu_N1-P        | FAM-ACCOCGCAATTGGTTTGTTGGAC-     |           |               |
|                 |                    | CDC-BHQ1                          |           |               |
| Spike protein   | CRM_S-F            | GACATACCATGTGGTGAG              | 83        | (JRC 2020)    |
|                 | CRM_S-R            | TGACTAGCTCAGCTGAGG              |           |               |
|                 | CRM_S-P            | FAM-AGAATCGCFAATTCCTTCGGCGG-    |           |               |
|                 |                    | CDC-BHQ1                          |           |               |
|                 | RdRp_SARSr-F       | GTGARATGGTCATGTTGGGCGG           | 100       | (Corman et al. 2020) |
|                 | RdRp_SARSr-R       | CARATGTAAASACACTATGACATA         |           |               |
|                 | RdRp_SARSr-Pv2019  | TAMRA-CAGTGTGAGCCCTCACTAGGAGAATGC-BHQ2 | |               |
|                 | RdRp_SARSr-Pv1-general* | HEX-CCAGGTGGWACRTCATCGMGGTGATGC-BHQ1 | |               |

* these were not added from March 2021 onwards
experimental runs. The LOD was estimated by LOD<sub>50</sub> (Marien et al. 2017); detecting 5 copies of N1 in the reaction, i.e. 625 copies in 1 L of wastewater. For the determination of the LOD<sub>95%</sub> it would be necessary to perform follow-up analyzes as described in or JRC Technical report on the verification of analytical methods for GMO testing (Hougs et al. 2017).

2.7. Correlation analysis

A generally accepted model correlating the number of COVID-19 cases in population and the concentration of SARS-CoV-2 RNA in wastewater has not been established so far. Moreover, the catchment areas selected in this study are extremely diverse (size, type of sewers, type of typical activities). Therefore, we applied a simple linear fit to describe the correlation between log-transformed number of positive cases and log-transformed data for number of N1 gene copies per liter of wastewater. As the log-transformed data were normally distributed (data not shown), the least square method could be used to fit the data. Person correlations with p-values greater than 0.05 (probability level α = 0.05) were considered statistically significant.

Due to technical reasons (development of sampling network in Prague, availability of funding), high amount of samples have been taken only during spring 2021 (April to May) when relatively few COVID-19 cases have been officially reported. This resulted in high number of samples below detection limit (54%). To consider this, only samples taken when more than 100 positive cases per hundred thousand inhabitants were officially reported in Prague (i.e. between 12 September 2020 and 5 May 2021) were included in the correlation analysis. In this period, the proportion of the samples below the limit of detection was less than 25 % and the occurrence of negative samples only weakly correlated (R<sup>2</sup> = 0.395) with the epidemic situation (data not shown). It can be assumed that the occurrence of negative samples was due to issues related to the experimental procedure (mainly the presence of PCR-inhibitory agents) without any relation to the actual amount of viral RNA in the sample. Therefore, the negative samples were not taken into account for the correlation analysis.

3. Results and discussion

3.1. The course of the COVID-19 epidemic in Prague

After successfully controlling the first wave of the COVID-19 epidemic in spring 2020 by variety of strict measures (e.g. mandatory to cover mouth and nose while being in public space; all restaurants, pubs, bars, etc. closed; non-essential shops closed; distance teaching of students in all types of schools; gatherings in public space limited 2, 6, or 10; social distancing at least 2 m; see (ISC, 2021 , Ministry of Health of the Czech Republic, 2021b) for more information), the Czech government eased the anti-COVID restrictions during summer 2020. This approach resulted in a massive increase in COVID-positive cases from October to November (a second wave) which was counteracted on 14 October by strict measures that included closing shops, restaurants, and other services, closing most of the schools and student dormitories, restricting free movement and public gatherings, etc. The third wave of new COVID-19 cases in December 2020 and January 2021 followed a partial easing of anti-epidemic measures in December (opening shops). The fourth wave started in February 2021 (probably resulting from the spread of new, more infectious variants of the SARS-CoV-2 virus). The situation in Prague was very similar to the rest of the Czech Republic, and the maximum numbers of newly reported cases per day were as high as 1,708, 2,065, and 2,037 in the second, third, and fourth wave, respectively (Ministry of Health of the Czech Republic, 2021a). The maximum number of active COVID-19 cases reached 1,506 per hundred thousand inhabitants on 9 March (Fig. 2).

3.2. Correlation of N1 gene concentration in sewage with epidemic data based on clinical testing

3.2.1. The detection of SARS-CoV-2 at small sewers in the long-term sampling

The presence of SARS-CoV-2 RNA was detected at all locations of Prague sewer network chosen for this study. However, the frequency of positive samples, N1 gene concentrations, and the correlation between N1 gene concentration and the number of positive COVID-19 cases were dramatically different between different locations (Fig. 3). Overall, 46% of all tested samples were positive, 47% were negative and 7% could not be evaluated due to PCR inhibition. The relatively large portion of
negative samples is given by the fact that a disproportionally high number of the samples were collected during the period of the declining epidemic (22 March–23 May 2021, Fig. 2). The numbers of N1 gene copies in the positive samples ranged from $10^{1}$ to $10^{7}$ per L of wastewater.

3.2.2. The correlation at small sewers in the long-term sampling

To demonstrate the value of the data measured in wastewater for the monitoring of COVID-19 epidemics, N1 gene copy numbers per L of wastewater on a logarithmic scale were correlated to the estimated active cases in Prague based on clinical testing (Fig. 4). The highest correlation between N1 concentration in wastewater and clinical epidemic data was observed at Locations 3, 5, 7, and 10 (determination factor $R^2$ higher than 0.5). Of these 4 locations, 3 are dominantly residential areas with more than 7000 registered inhabitants and one includes a shopping mall. As the shopping mall (Location 10) involved a grocery shop that was open throughout the pandemic, it can be assumed that the shopping mall has been used by large amount of people even during the periods with strict epidemic restrictions.

As shown in Fig. 5, the correlation was poor ($R^2$ less than 0.3) in all other small locations, i.e. non-residential areas (Locations 8, 9, and 11) and residential areas with less than 7000 inhabitants (Location 1, 2, 4, and 6).

A surprisingly good agreement was observed among the trend lines obtained for all residential areas larger than 7,000 inhabitants and the shopping mall (Fig. 4A, Table 3), showing that the correlation equations can be extrapolated from one larger (> 7,000 inhabitants) residential area to another. This finding, however, cannot be without further research applied to other places, e.g. to different cities.

It was not clear whether a close correlation between the copy number of N1 gene detected in wastewater and epidemic data should be expected for all monitored locations. E.g. Mota et al. (2021) suggested that sub-sewersheds may only be used to identify epidemic hotspots, but not for the monitoring of the general epidemic situation. Besides the fact

---

Fig. 3. SARS-CoV-2 N1 gene concentration in wastewater (logarithmic scale for better visualization) and the number of registered positive cases of COVID-19 as reported by the Czech Ministry of health at selected locations monitored since September 2020. Full black line – positive cases in Prague per 100 thousand inhabitants estimated from clinical testing; Dashed grey line – positive cases at respective locations per 100 thousand inhabitants estimated from clinical testing; Filled black points – decimal logarithm of N1 gene copy number per liter of wastewater; Open circles – negative samples (N1 gene under detection limit).
that the real number of infected individuals active at the monitored locations is unknown, the effect of time must also be taken into account when interpreting the data. On some occasions, wastewater-based data exhibited a 1 to 2 weeks lead (Location 11 – student dormitories – in September) or delay in epidemic progression compared to the epidemic data (Location 6 – residential area – in September 2020, Fig. 3). Nonetheless, our data show relatively good correlation with the clinical data for all residential areas with more than 7,000 registered inhabitants (Fig. 5a, Table 3).

3.2.3. Predictive value of SARS-CoV-2 RNA monitoring in wastewater at small locations

Some studies have highlighted the predictive value of SARS-CoV-2 RNA monitoring in wastewater, claiming that the increase in copy numbers in wastewater (Ahmed et al. 2020, Haramoto et al. 2020, Sherchan et al. 2020, Wu et al. 2020a) or sewage sludge (Peccia et al. 2020) precedes the rise of clinically detected COVID-19 cases by up to two weeks. Medema et al. (2020) detected SARS-CoV-2 RNA in sewage from Amersfoort (the Netherlands) 6 days before the first clinical cases were reported, and La Rosa et al. (2021) showed that the virus was circulating in Italy even in December 2019. Similarly, during the SARS epidemic in 2002, Wang et al. (2005) detected coronavirus RNA originating from the feces of patients hospitalized with SARS-CoV-2 infection in sewage up to 8 days prior to the outbreak of the epidemic.

This predictive value was not consistently demonstrated in our research as the correlation between number of COVID-19 cases and the number of N1 gene copies in wastewater (R² value) did not improve while virtually shifting the wastewater-based data up to 2 weeks ahead or back compared to the number of positive case (shown for Location 7 in Fig. 6F). The reason for this observation can be the low frequency of sampling, which was chosen in this study for capacity reasons.

3.2.4. The correlation observed in the main Prague sewers during the intensive sampling campaign

The trunk sewers, representing large regions of Prague, were only monitored during the intensive sampling campaign (22 March – 21 May 2021, Fig. 2), when the COVID-19 epidemic have been declining already. However, the correlation between N1 gene copy numbers and the number of active COVID-19 cases in Prague (Fig. 4b) was generally better (R² between 0.64 and 0.71) than the best-correlating residential areas with the exception of trunk sewer F (S3). No systematic influence on the correlation coefficient was observed for the size of population served by the trunk sewers, which ranged from 51 to 671 thousand inhabitants (Fig. 5b). Again, the trend lines obtained for individual main sewers were very similar to each other (Fig. 4b), and the calibration performed with the data from Sewer ACK (S1) was successfully used to estimate the number of active COVID-19 cases for all the other main sewers (Fig. 6E).

3.2.5. The importance of wastewater flow measurement

The representativeness of the concentration-based data collected in all trunk sewers (ACK, C, K, and F) was compared to the data normalized to flow and population (copy number per person per hour). This approach is often used for the correlation with clinical-based data (Huisman et al. 2021). However, in our study were the determination factors almost identical for these two methods at all trunk sewers (Figs. 4C and 5B). This shows that concentration-based data were as valid as the flow-normalized data. Interestingly, the agreement between the flow-normalized data obtained for individual trunk sewers (Fig. 4B and C, Table 3) was much worse than that of concentration-based data (Fig. 4B). This may be given e.g. by inconsistency between the official and actual number of people living at certain parts of the city.

As flow measurement was not available at the small locations, the measurements of SARS-CoV-2 RNA concentration was used for the estimation of COVID-19 cases and these data were usable for epidemic monitoring without using complex mathematical models such as the one
3.2.6. Predictive value of SARS-CoV-2 RNA monitoring in wastewater in trunk sewers

The data for the trunk sewers came almost exclusively from the period of epidemic decline, when the wastewater-based data were biased by the long excretion of viral RNA in stools, which might persist for up to 2 weeks after the patient is considered COVID-19-negative (Hong et al. 2021, Wu et al. 2020b). The decline in SARS-CoV-2 RNA occurrence in Prague wastewater slightly lagged (2 to 8 days) behind the decrease in positive samples in clinical data as shown by the increase of determination coefficient between wastewater-based and clinical-based data when this delay is applied (Fig. 6F). It should be noted that this delay of the wastewater-based data was confirmed only for the decline of the pandemic, as the monitoring of the trunk sewers did not cover the period of increasing incidence of the disease in the population.

3.2.7. Comparing data from small locations and trunk sewers

The averaged trend line obtained for larger residential areas (Locations 3, 5, 7, and 10) was steeper than the trend line for the main sewers (Fig. 6F). Besides other factors, this was given by the fact that at the respective locations that cover in total less than 2.5% of Prague’s population, there was a very low probability of the occurrence of positive cases during the periods with low COVID-19 incidence. Indeed, few positive samples were detected at Locations 1 to 22 when the number of positive cases in Prague was lower than 300 per 100 thousand inhabitants (Fig. 3). In contrast, few negative samples were detected in the trunk sewers when more than 100 COVID-19 cases were reported per 100 thousand inhabitants (Fig. 6). This, indeed, confirms that monitoring sewer sheds larger than approximately 50 thousand inhabitants are sensitive enough to detect one positive case in the population of one thousand individuals.

Interestingly, following the week of 2–8 May, when all samples from the main sewers were negative, the numbers of N1 gene copies in the samples from the main sewers started to grow again. Clinical data has not confirmed this trend by the time of writing this report.

3.3. Epidemiological value of wastewater-based data on SARS-CoV-2 occurrence in Prague wastewater

After the very recent paper by Mota et al. (2021) reporting decentralized monitoring of the sewers of Belo Horizonte (Brazil) for three months in 2020, this is the first European study comparing the epidemic relevance of data collected during 10 months at local neighborhoods of different sizes (approx. 100 to 14 000 registered residents) and areas covering large parts (main sewers S2, S3, and S4) or more than half of the city (the main sewer S1) of 1.32 mil. inhabitants. The correlation between wastewater-based and clinical testing-based data was surprisingly good for residential areas with more than 7000 registered inhabitants and was almost as good as the same correlation obtained for the main sewers (51,000 to 671,000 inhabitants). Moreover, the data obtained from the small residential areas seemed to be quicker than those from the trunk sewers, at least during the decline of the COVID-19 pandemic when the data from trunk sewers had 2 to 8 days delay.
Mota et al. (2021) showed that COVID-19 hotspots could be identified in the city based on data generated by decentralized sewage monitoring especially in vulnerable neighborhoods such as favelas, which tend to be densely populated and with limited sanitation infrastructure. Our study suggests on long-term data that epidemic hotspots escaping the attention of public authorities may occur also in a major European city where such vulnerable areas are not expected (Fig. 4B).

The decrease in N1 gene copy numbers in the samples collected at small locations (Locations 1 – 22) was generally not delayed compared to the clinical data (Figs. 3, 6F). However, positive samples were repeatedly identified at locations such as 5 (Family houses; Fig. 3), 19 (Hospital; data not shown), and 22 (Bus terminal; data not shown), even while the total numbers of positive cases in Prague were well below 100 per 100 thousand inhabitants. Public authorities have not recognized these hotspots, as numerous COVID-19 cases were not reported for the individuals officially registered at given locations (Fig. 3).

This observation indicates that problematic areas with the occurrence of infected individuals could be identified despite the low average epidemic numbers. As a result, targeted epidemic measures could be applied in these areas. The correlation between wastewater-based data and COVID-19 cases officially registered at the location monitored (Fig. 3) was assessed for each location separately (data not shown) and it was generally worse than the correlation with Prague-wide data. This is another hint that the spatial distribution of COVID-19 cases may be better described by monitoring wastewater than by simply analyzing the official addresses of COVID-19-positive individuals.

Huisman et al. (2021) showed the feasibility of Reproduction factor (Fig. 6E and F).

Mota et al. (2021) showed that COVID-19 hotspots could be identified in the city based on data generated by decentralized sewage monitoring especially in vulnerable neighborhoods such as favelas, which tend to be densely populated and with limited sanitation infrastructure. Our study suggests on long-term data that epidemic hotspots escaping the attention of public authorities may occur also in a major European city where such vulnerable areas are not expected (Fig. 4B).
Variations between individual sampling locations. 

Finally, we showed that even grab samples can be successfully used when 24h composite samples cannot be collected for practical reasons (lack of resources, badly accessible sewers etc.). It is important to mention that precise time-dependent sampling (typically during the morning peak) was kept and the effects of weather conditions were avoided (i.e. the samples were never collected shortly after major precipitation events). Indeed, also Black et al. (2021) showed that grab samples may be of high epidemiological value when the time of sampling is chosen appropriately.

3.4. The inhibition of RT-qPCR

Wastewater is a complex matrix that contains several potential inhibitors of reverse transcription as well as PCR. Such inhibitors include humic acids, fulvic acids, humic material, metal ions, polyphenols from the environment, and complex polysaccharides, bile salts, lipids, and urate from stools (Schrader et al., 2012). These inhibitors of the molecular biology assay should be removed during sample preparation for RT and PCR, mainly during RNA isolation.

We used the TRIzol™ isolation method, as it is reliable and relatively inexpensive. Other studies also recommended this method for municipal wastewater samples. E.g. Torri et al. (2021) observed the highest recovery of phage RNA while using pre-concentration by PEG followed by TRIzol™ RNA isolation. Even though the RT-qPCR was inhibited in some samples collected in our study, this was relatively rare and did not jeopardize the interpretation of our data.

4. Conclusions

This paper reports the presence of SARS-CoV-2 RNA in wastewater collected at different locations such as residential areas of various sizes, hospitals, city center areas, student dormitories, transportation hubs, and commercial areas within a large city (Prague, Czechia). The data obtained at the main Prague sewers were more consistent than those obtained from the small sewers. However, the correlation between wastewater-based data and data from clinical testing was good ($R^2 > 0.5$) for the residential areas with more than 7,000 registered inhabitants. The latter was true despite the fact that only grab samples were collected from the small sewers. This study also shows that monitoring SARS-CoV-2 RNA in wastewater sampled from small sewers can identify an isolated occurrence of COVID-19-positive cases in local neighborhoods. This can be highly valuable while tracking COVID-19 hotspots within a large city (Fig. 7).

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.

Acknowledgments

This work was financially supported by Technology Agency of the Czech Republic (TA CR) grant no. SS01020112 “Technologies for removal of antibiotic resistance genes from sewage sludge applied in agriculture”. The authors also wish to thank PKV a.s., VZU, and UCT Prague for their generous financial and material support. The valuable assistance of Karel Behounek (Aqua-Contact, v.o.s.) and Irena Novakova (Prague Airport) during wastewater sampling is greatly appreciated, as well as the dedicated lab work of Nelly Matouskova (Czech Health Institute), Marco Lopez, Sara Doubovská, Katarina Hanusova, and Dominik Tomanek (students of UCT Prague). Finally, we thank to Christof Uisk (student of UCT Prague) for creating the maps of the catchment areas of all sewers.

References

Acosta, N., Bautista, M.A., Hollman, J., McCalder, J., Beaudet, A.B., Man, L., Waddell, B.J., Chen, J., Li, C., Kuzma, D., Bhatnagar, S., Leal, J., Meddings, J., Hu, J., Cabal, J.J., Ruecker, N.J., Naugler, C., Pillai, D.R., Achari, G., Ryan, M.C., Conly, J.M., Frankowski, K., Hubert, C.R., Parkins, M.D., 2021. A multicenter study investigating SARS-CoV-2 in tertiary-care hospital wastewater: viral burden correlates with increasing hospitalized cases as well as hospital-associated transmissions and outbreaks. Water Res. 201, 117369.

Ahmed, F., Islam, M.A., Kumar, M., Hosain, M., Bhattacharya, P., Islam, M.T., Hossen, F., Hossain, M.S., Islam, M.S., Uddin, M.M., Islam, M.N., Bahadur, N.M., Didar-ul-Alam, M., Reza, H.M., Jakaria, M., 2021a. First detection of SARS-CoV-2 genetic variant in the vicinity of COVID-19 isolation Centre in Bangladesh: variation along the sewer network. Sci. Total Environ. 776, medRxiv.

Ahmed, W., Angel, N., Edson, J., Bivins, K., Bibby, K., Bivins, A., O’Brien, J.W., Choi, P.M., Kitijama, M., Simpson, S.L., Sherman, P., Thomas, K.V., Mueller, J.F., 2020. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. Sci. Total Environ. 728, 138764.

Ahmed, W., Tschirke, B., Bertsch, P.M., Bibby, K., Bibby, A., Choi, P., Clarke, L., Dwyer, J., Edson, J., Nguyen, T.M.H., O’Brien, J.W., Simpson, S.L., Sherman, P., Thomas, K.V., Verhagen, R., Zaugg, J.L., Mueller, J.F., 2021b. SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission in the community: a temporal case study. Sci. Total Environ. 761, 144216.

Baldovin, T., Amoroso, I., Fonzo, M., Buja, A., Baldo, V., Coccio, S., Bertoncino, C., 2021. SARS-CoV-2 RNA detection and persistence in wastewater samples: An experimental network for COVID-19 environmental surveillance in Padua, Veneto Region (NE Italy). Sci. Total Environ. 760, 143329.

Black, J., Aung, P., Nolan, M., Roney, E., Poon, R., Hennessey, D., Crobie, N.D., Deere, D., Jex, A.R., John, N., Barker, L., Scales, P.J., Usher, S.P., McCarty, D.T., Schang, C., Schmidt, J., Myers, S., Begue, N., Kauzner, C., Thoryl. B., Duce, J., Moris, P., Lau, M., Sarkis, S., 2021. Epidemiological evaluation of sewage surveillance as a tool to detect the presence of COVID-19 cases in a low case load setting. Sci. Total Environ. 786, 147469.

CDC (2020) Division of Viral Diseases, Centers for Disease Control and Prevention, http://www.cdc.gov/coronavirus/2019-ncov/downloads/rt-pcr-panel-primer-probes.pdf. Accessed on 2020 30 June.

Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K., Bleicker, T., Brunk, S., Schoedter, J., Schmidt, M.L., Mulders, D.G., Haagmans, B.L., van der Veer, B., van den Brink, S., Wijmenga, L., Gersdorff, G., Romette, J.J., Ellis, J., Zambron, M., Peiris, M., Goossens, H., Reusken, C., Koopmans, M.P., Drosten, C., 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 25 (3). 2000045.

Crank, K., Chen, W., Bivins, A., Lowry, S., Bibby, K., 2022. Contribution of SARS-CoV-2 RNA shedding routes to RNA loads in wastewater. Sci. Total Environ. 806, 150376.

Didar-ul-Alam, M., Reza, H.M., Jakariya, M., 2021a. First detection of SARS-CoV-2 genetic variant in the vicinity of COVID-19 isolation Centre in Bangladesh: variation along the sewer network. Sci. Total Environ. 776, medRxiv.

Haramoto, E., Malla, B., Thakali, O., Kitajima, M., 2020. First environmental surveillance from wastewater can be done in real time, is robust, rapid, low cost, and applicable for various wastewater matrices. Water Res. 201, 117369.
