BRIEF REPORT

environmental
microbiology reports   sfam

SARS-CoV-2 whole-proteome sequences from environment as an indicator of community viral distribution, evolution and epidemiological dynamics: A cohort analysis of Austria

Murali Aadhitya Magateshvaren Saras | L. Ponoop Prasad Patro |
Patil Pranita Uttamrao | Thenmalarchelvi Rathinavelan [ID]

Department of Biotechnology, Indian Institute
of Technology Hyderabad, Kandi, Telangana,
India

Correspondence
Department of Biotechnology, Indian Institute
of Technology Hyderabad, Kandi, Telangana
502285, India.
Email: tr@bt.iith.ac.in

Abstract

Several investigations have been carried out to detect SARS-CoV-2 samples from the environment such as sewage waters and surface swabs. Whole-proteome sequence analysis of 847 SARS-CoV-2 genome sequences collected from the environment in Austria during 2021 and deposited in GISAID indicates that alpha and delta are two dominant variants, coinciding with the human clinical samples with a Pearson correlation coefficient in the range of 0.58 (alpha variant) to 0.82 (delta variant). Both environmental and human samples show that Austrian SARS-CoV-2 alpha variant is found to possess N protein R203K and G204R/P mutations, whereas they are absent in the delta variant. SARS-CoV-2 delta variant is continuously seen in both the environmental and human clinical samples from the month of September 2021 and it spiked in November 2021, which is directly reflected in the increase of the number of SARS-CoV-2 infections and deaths in Austria during November 2021. Thus, the results presented here indicate that the environmental SARS-CoV-2 whole-genome sequences collected from Austria reflect the community viral distribution, evolution and the concomitant epidemiological dynamics. Since SARS-CoV-2 keeps evolving, the results presented here further suggest the need to monitor the environment for the early detection of SARS-CoV-2 variants to take appropriate precautionary measures.

INTRODUCTION

SARS-CoV-2 is the etiological agent behind the ongoing pandemic. Several studies have detected SARS-CoV-2 RNA in environment such as water, soil, wastewater, door knobs and taps due to the human shedding of the virus (Patel et al., 2021; World Health Organization, 2020). For instance, traces of SARS-CoV-2 RNA have been found via wastewater surveillance in the USA (Sherchan et al., 2020), Italy (La Rosa et al., 2020), Spain (Randazzo et al., 2020) and the Netherlands (Gertjan Medema et al., 2020). There is an evidence that suggests wastewater monitoring could facilitate the prediction of SARS-CoV-2 community transmission and prevent the viral outbreak by taking proper countermeasures (Ahmed et al., 2021). Thus, SARS-CoV-2 environmental surveillance can

act as a valuable early warning tool for tracking the community circulation of SARS-CoV-2 as it represents the presence of virus irrespective of symptomatic, asymptomatic, pre-symptomatic and misdiagnosed carriers (Chavarria-Miro et al., 2021). Furthermore, environmental surveillance has shown not only to provide real-time and unbiased disease surveillance but also to be instrumental in identifying the local community spread and abundance of SARS-CoV-2 lineages that are circulating in the community (CDC, 2021; European Commission: DG Environment, 2021; Wei Lin Lee et al., 2021). For instance, dissemination of the contagious SARS-CoV-2 alpha variant (PANGO lineage is B.1.1.7) in Missouri (USA) wastewater has been detected much ahead of the first clinical case reported (Anthes, 2021). To this end, a comparative analysis of Austrian environmental and

human SARS-CoV-2 protein sequences translated *in silico* from the whole-genome sequence data collected during 2021 has been carried out to explore the use of environmental surveillance in tracking the viral spread and the concomitant community epidemiological dynamics.

RESULTS AND DISCUSSION

The public health emergency created by SARS-CoV-2 pandemic necessitates the effective surveillance system to mitigate the community spread by taking appropriate preventive measures. This is particularly important in the scenario of frequent emergence of several variants of concerns (Choi & Smith, 2021; Janik et al., 2021; Thye et al., 2021). One such worrisome example is the recently emerged highly transmissible SARS-CoV-2 omicron variant against which the vaccines are less effective (Alexander Wilhelm et al., 2021; Frederik Plesner Lyngse et al., 2021; Syed et al., 2022). Previous studies have shown that wastewater-based disease monitoring of SARS-CoV-2 is effective in quickly detecting the local viral spread (Kitajima et al., 2020). Thus, to explore the benefit of using environment as an effective surveillance tool, a comparative whole phyloproteogram analysis of SARS-CoV-2 sequences obtained from the environment and the humans in Austria has been performed here (refer supplementary file S1 for Materials and methods).

Fifty-two high and 80 moderate recurring mutations are found in SARS-CoV-2 sequences from Austrian clinical samples

Since the amino acid mutations may provide direct inference about the changes in viral genome on antigenicity, immunogenicity, transmissibility and pathogenicity, the amino acid mutations of SARS-CoV-2 are analysed. There are 52 high and 80 moderate recurring mutations in the SARS-CoV-2 sequences collected from Austrian human sequences (Supplementary Table S1). Although fifty HR mutations are found in the environmental sequences, only six MR mutations are

present in the environment (Supplementary Table S2). This perhaps is due to the smaller sample size of the environmental sequences. A detailed inspection indicates that Austrian SARS-CoV-2 from environmental and humans have Spike protein mutations corresponding to two variants of concern, namely, alpha and delta. Furthermore, SARS-CoV-2 alpha variant sequences are found to invariably contain N protein:R203K, N protein:G204R/P, Nsp6 deletions: S106-, G107-, F108-, Nsp3:A890D and Spike:S982A both in human and environmental sequences. Notably, delta variant does not have these mutations, instead, it consistently possesses Spike:L452R, Spike:T19R, Spike:T478K, Nsp3: P1469S, ORF3aprotein:S26L, Spike:P681R, M protein: I82T and N protein:D377Y. Besides, high recurring Spike: N501Y, N protein:S235F, Nsp3:T183I, ORF8protein: Y73C, Spike:V70-, Spike:A570D, Spike:H69-, Spike: D1118H, Nsp3:I1412T, Spike:P681H and Spike:Y145-proteome mutations are seen to be specific to alpha variant, whereas high recurring Nsp12:G671S, N protein: R203M, ORF7aprotein:V82A, Nsp13:P77L, ORF7aprotein:T120I, Spike:E156G, Spike:F157-, Spike:R158-, Nsp3:P1469S and Nsp4:T492I mutations are specific to delta variant (Supplementary Table S3).
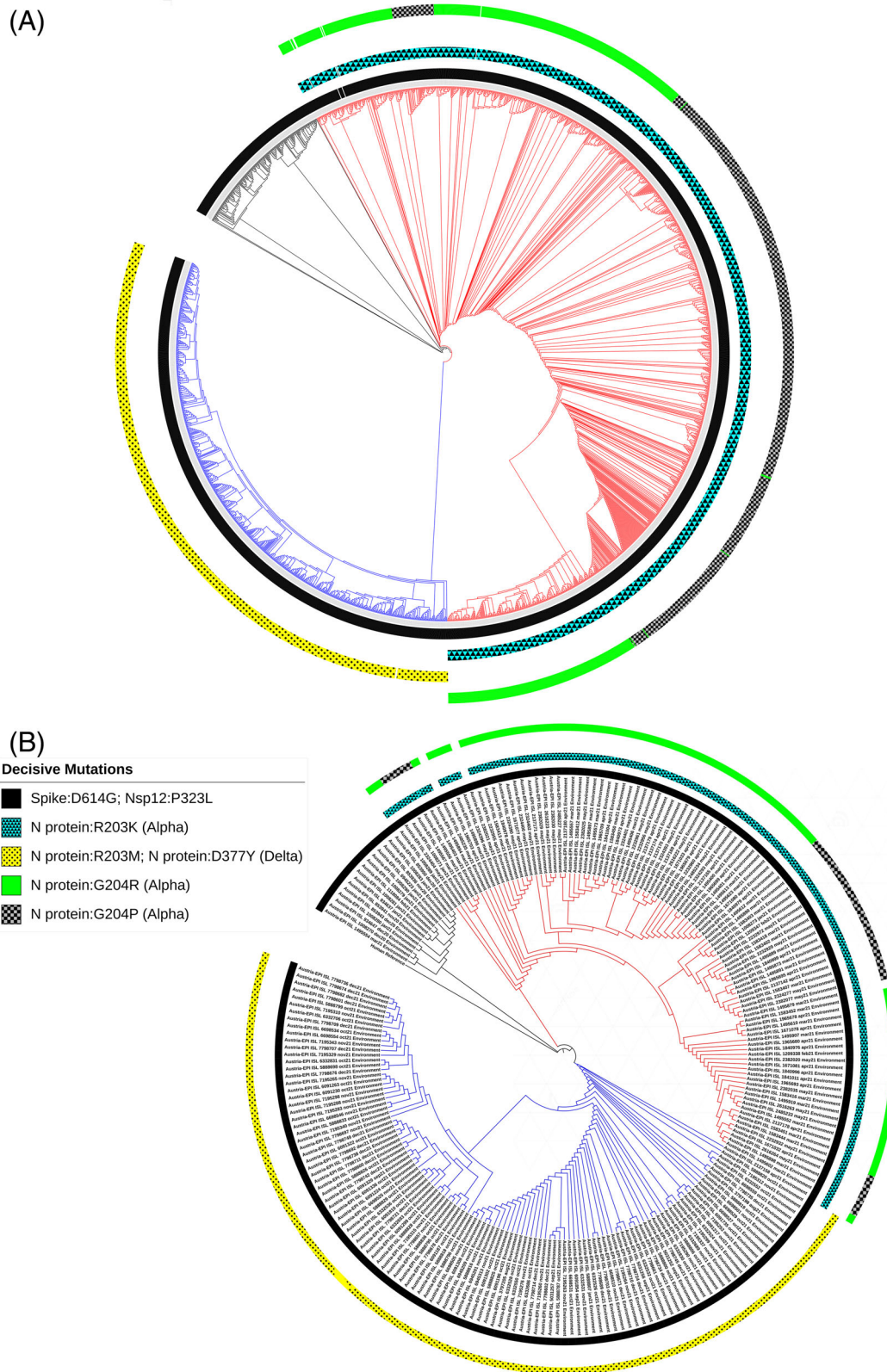
Dominance of alpha variant during the first half of 2021 and delta variant during the second half of 2021

The month-wise distribution of alpha (lineage B.1.1.7) and delta (lineage B.1.617.2 and its sublineages) variants [Figure 1(A)] shows that the alpha variant is dominant during the months of March–April 2021 in human, and it declines after August 2021. Nonetheless, delta variant dominates beyond September 2021. Very importantly, SARS-CoV-2 environmental sequence also reflects the same trend [Figure 1(B)]. Notably, the first occurrence of SARS-CoV-2 delta variant reported in Austria during January 2021 is from an environmental sample. However, it becomes dominant soon after the decline of alpha variant during June 2021.

A thorough analysis indicates that many subvariants of alpha and delta found in the human samples
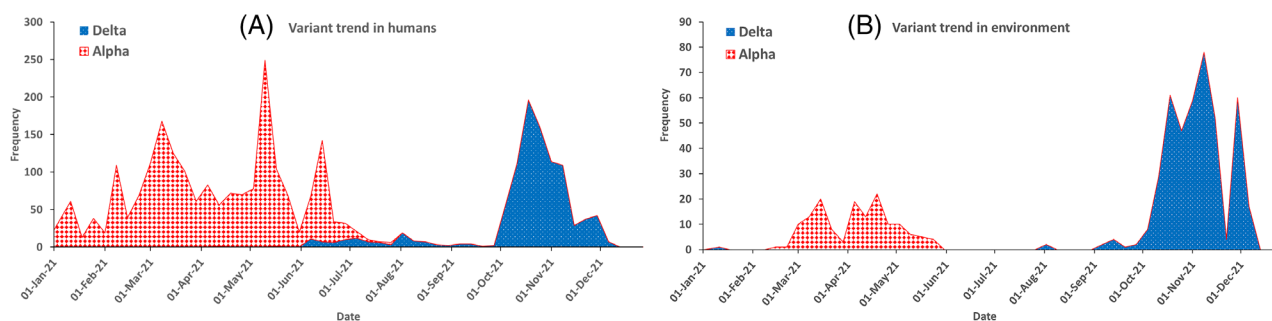
FIGURE 1   Graph representing the month-wise distribution of alpha (patterned, red) and delta (filled, blue) variants in Austria human clinical (A) and environmental (B) samples during the year 2021. Note that SARS-CoV-2 sequences are represented on weekly basis

FIGURE 2 Phyloproteogram illustrating the dominance of alpha (coloured blue) and delta (coloured yellow) variants in human clinical samples (A) and environment samples (B) in Austria during the year 2021. Note that the black branches do not belong to any variants of concern

are also seen in the environmental samples. For instance, environmental as well as human sequences have been identified with mutations such as N protein: G204R/P (99.91%), Spike:P681H (99.58%), Spike:

T716I/V (99.90%) and N protein:D3L/E (94.70%) associated to th alpha variant (B.1.1.7) (Supplementary table S3). Coincidentally, AY.43 lineage of delta variant is found with the highest percentage (8.87%) in the
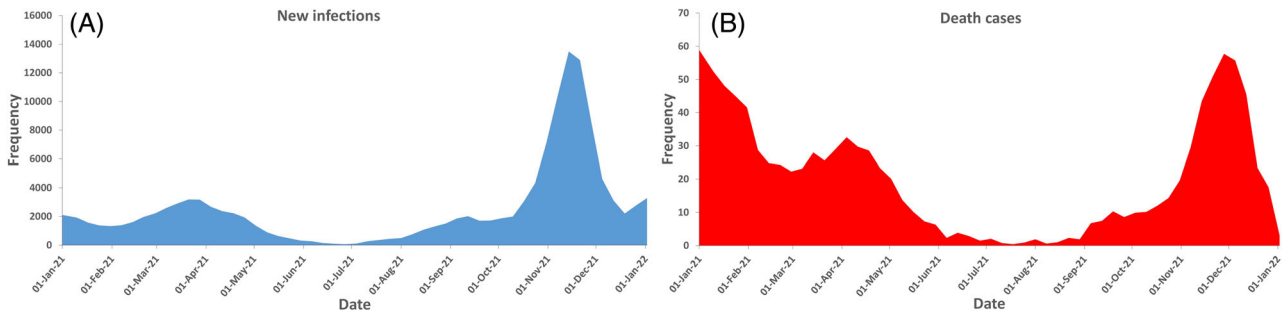
FIGURE 3 Graph depicting the daily new infections (A) and deaths (B) reported in Austria during January 2021 to December 2021. Note that the values are reported weekly. Source – ourworldindata.org (https://ourworldindata.org/coronavirus/country/austria)

human sequences. AY.122 (5.12%), B.1.617.2 (3.81%), AY.4 (2.05%) and AY.46.6 (2.26%) lineages of the delta variant constitute a total of 28.55% amongst all the cases reported in humans and are found to have key mutations N protein:Q9L (24.18%), Spike:T95I (25.32%), Nsp2:K81N (12.13%) and ORF7aprotein: P45L (10.70%). These sub-variants are also seen in the environmental samples (Supplementary Table S4 and S3). The positive Pearson coefficients estimated for alpha (0.58 with a *p*-value <0.01) and delta (0.82 with a *p*-value <0.01) variants further indicate the close correlation between the human and environmental SARS-CoV-2 samples.

Whole phyloproteogram of SARS-CoV-2 environmental sequence reflects the viral evolution in SARS-CoV-2 clinical samples

To further understand the viral evolution in Austria, whole phyloproteogram of SARS-CoV-2 environmental and human sequences are analysed independently. The phyloproteomic analysis has been carried out to see the emergence of sub-variants from an ancestral variant. It also clearly would provide information about the descendent variant. All the lineages in the phylopro-teogram invariably have Spike:D614G and Nsp12: P323L mutations. As described previously in the muta-tion analysis, two major clades are found in the whole phyloproteogram which correspond to the alpha and delta variants (Figure 3). The alpha variant clade invari-ably has the N protein R203K and G204R/P mutations. Further divergence seen in this clade is associated with the mutations in other proteins: Spike:N501Y, N pro-tein:S235F, Nsp3:T183I, ORF8protein:Y73C, Spike: V70-, Spike:A570D, Spike:H69-, Spike:D1118H, Nsp3: I1412T, Spike:P681H and Spike:Y145-. Similarly, the delta variant clade has high recurring Nsp12:G671S, N protein:R203M, ORF7aprotein:V82A, Nsp13:P77L, ORF7aprotein:T120I, Spike:E156G, Spike:F157-, Spike:R158-, Nsp3:P1469S and Nsp4:T492I, in addi-tion to the other mutations that define various sub-clades (viz., sub-lineages or sub-variants of delta). The

human whole phyloproteogram also has the same trend, indicating that the amino acid variations seen in SARS-CoV-2 environment sequences follow the same trend as the clinical samples.

Furthermore, B.1.1.7 PANGO lineage corresponding to the alpha variant is found both in the environmental and human SARS-CoV-2 samples. Forty-five PANGO sub-lineages that are associated with delta variant are found in the human sequences. Among them, 27 are found in the environmental SARS-CoV-2 sequences. Additionally, one PANGO sub-lineage is found in the envi-ronmental sample, but is absent in the human samples. It is to be noted that the 27 PANGO sub-lineages which are seen in the environmental SARS-CoV-2 sequences occur at a higher frequency in the humans compared to the rest of the PANGO sub-lineages.

Thus, the mutation (discussed above) and variant analyses (Figure 1) along with the phyloproteograms (Figure 2) of SARS-CoV-2 sequences obtained from clini-cal and environmental samples clearly pinpoint that alpha and delta variants are the dominant variants found in Austria. The other variants such as Zeta, Iota, Gamma and Beta variants are identified in negligible percentage frequencies (0.02%, 0.02%, 0.15% and 0.22%).

Most importantly, the month-wise infection caused by alpha and delta variants in humans is also reflected in the environmental sequences (Figure 1). Both show that while the alpha variant has peaked during March–June 2021, the delta variant has dominated during October–December 2021. Strikingly, the months of occurrence of these variants coincide with the second and third waves of SARS-CoV-2 in Austria as can be witnessed from the number of infections [Figure 3(A)] and deaths [Figure 3(B)]. Thus, environment samples reflect the spread of different variants in the local com-munities and the associated disease epidemiological dynamics in Austria during the year 2021.

CONCLUSIONS

To assess the potential of environmental surveillance for the detection of SARS-CoV-2 variants circulating in

the local community, a detailed analysis of SARS-CoV-2 sequences collected from environment samples from Austria has been carried out and the results are compared with SARS-CoV-2 clinical samples. The results clearly pinpoint that the environment samples exhibit a similar trend as the clinical samples by reflecting the dominance of SARS-CoV-2 alpha and delta in Austria at two different periods during 2021. Furthermore, Pearson correlation coefficient exhibits a moderate (0.58 in the case of alpha variant) to high (0.82 in the case of delta variant) positive correlation between environmental and clinical sequences. The environmental samples indicate the circulation of several sub-variants of alpha and delta in the community. Most importantly, the alpha and delta variants in environmental samples reflect the second and third waves of SARS-CoV-2 in Austria. Thus, the results indicate that the environment-based disease monitoring would be effective in the surveillance of SARS-CoV-2 as it provides accurate information about the local spread of variants by representing the presence of virus irrespective of symptomatic, asymptomatic, pre-symptomatic and misdiagnosed carriers. As the world is witnessing the frequent emergence of several variants of concerns of SARS-CoV-2 such as delta, omicron, and their sub-variants, environmental surveillance may be helpful in taking appropriate precautionary measures as well as understanding the epidemiological dynamics of the disease. The result further suggests that environmental surveillance can act as a time- and cost-effective measure in countries which have lower screening capacity.

AUTHOR CONTRIBUTIONS

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT
The data used in the current study is provided as the supplementary material.

ORCID
Thenmalarchelvi Rathinavelan https://orcid.org/0000-0002-1142-0583

REFERENCES

Ahmed, W., Tscharke, B., Bertsch, P.M., Bibby, K., Bivins, A., Choi, P. et al. (2021) SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission in the community: a temporal case study. *Science of the Total Environment*, 761, 144216.

Alexander Wilhelm, M.W., Grikscheit, K., Toptan, T., Schenk, B., Pallas, C., Metzler, M. et al. (2021) Reduced neutralization of SARS-CoV-2 omicron variant by vaccine sera and monoclonal antibodies. *medRxiv*. https://doi.org/10.1101/2021.12.07.21267432

Anthes, E. (2021) From the wastewater drain, solid pandemic data. *The New York Times*.

CDC. (2021) *Wastewater surveillance*. Available at: https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html [Accessed 5th January 2022].

Chavarria-Miro, G. et al. (2021) Time evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in wastewater during the first pandemic wave of COVID-19 in the metropolitan area of Barcelona, Spain. *Applied and Environmental Microbiology*, 87(7), e02750-20.

Choi, J.Y. & Smith, D.M. (2021) SARS-CoV-2 variants of concern. *Yonsei Medical Journal*, 62(11), 961–968.

European Commission: DG Environment. (2021) Commission recommendation on a common approach to establish a systematic surveillance of SARS-CoV-2 and its variants in wastewaters in the EU. In: The European Commission (Ed.).

Frederik Plesner Lyngse, L.H.M., Denwood, M.J., Christiansen, L.E., Møller, C.H., Skov, R.L., Spiess, K. et al. (2021) SARS-CoV-2 omicron VOC transmission in Danish households. *medRxiv*. https://doi.org/10.1101/2021.12.27.21268278

Gertjan Medema, L.H., Elsinga, G., Italiaander, R. & Brouwer, A. (2020) Presence of SARS-coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in The Netherlands. *Environmental Science & Technology Letters*, 7(7), 511–516.

Janik, E., Niemcewicz, M., Podogrocki, M., Majsterek, I. & Bijak, M. (2021) The emerging concern and interest SARS-CoV-2 variants. *Pathogens*, 10(6), 633.

Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C.P., Hamilton, K.A. et al. (2020) SARS-CoV-2 in wastewater: state of the knowledge and research needs. *Science of the Total Environment*, 739, 139076.

La Rosa, G. et al. (2020) First detection of SARS-CoV-2 in untreated wastewaters in Italy. *Science of the Total Environment*, 736, 139652.

Patel, M., Chaubey, A.K., Pittman, C.U., Jr., Mlsna, T. & Mohan, D. (2021) Coronavirus (SARS-CoV-2) in the environment: occurrence, persistence, analysis in aquatic systems and possible management. *Science of the Total Environment*, 765, 142698.

Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A. & Sánchez, G. (2020) SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water Research*, 181, 115942.

Sherchan, S.P., Shahin, S., Ward, L.M., Tandukar, S., Aw, T.G., Schmitz, B. et al. (2020) First detection of SARS-CoV-2 RNA in wastewater in North America: a study in Louisiana, USA. *Science of the Total Environment*, 743, 140621.

Syed, A.M., Ciling, A., Khalid, M.M., Sreekumar, B., Chen, P.-Y., Kumar, G.R. et al. (2022) Omicron mutations enhance infectivity and reduce antibody neutralization of SARS-CoV-2 virus-like particles. *medRxiv*. https://doi.org/10.1101/2021.12.20.21268048

Thye, A.Y., Law, J.W.-F., Pusparajah, P., Letchumanan, V., Chan, K.-G. & Lee, L.-H. (2021) Emerging SARS-CoV-2 variants of concern (VOCs): an impending global crisis. *Biomedicine*, 9(10), 1303.

Wei Lin Lee, M.I., Armas, F., McElroy, K.A., Gu, X., Duvallet, C., Chandra, F. et al. (2021) Quantitative SARS-CoV-2 alpha variant B.1.1.7 tracking in wastewater by allele-specific RT-qPCR. *Environmental Science & Technology Letters*, 8(8), 675–682.

World Health Organization. (2020) *Status of environmental surveillance for SARS-CoV-2 virus*. Available from: https://www.who.int/news-room/commentaries/detail/status-of-environmental-surveillance-for-sars-cov-2-virus.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.