



Simultaneous detection and characterization of common respiratory pathogens in wastewater through genomic sequencing

Rachel C. Williams^{a,*}, Kata Farkas^a, Alvaro Garcia-Delgado^a, Latifah Adwan^a,
Jessica L. Kevill^a, Gareth Cross^b, Andrew J. Weightman^c, Davey L. Jones^{a,d}

^a School of Environmental and Natural Sciences, Bangor University, Bangor, Gwynedd LL57 2UW, UK

^b Science Evidence Advice Division, Health and Social Services Group, Welsh Government, Cathays Park, Cardiff, CF10 3NQ, UK

^c School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK

^d Food Futures Institute, Murdoch University, Murdoch WA 6150, Australia

ARTICLE INFO

Keywords:

COVID-19

Wastewater-based epidemiology

Respiratory viruses

Pathogens

Wastewater sequencing

Public health surveillance

ABSTRACT

Genomic surveillance of SARS-CoV-2 has given insight into the evolution and epidemiology of the virus and its variant lineages during the COVID-19 pandemic. Expanding this approach to include a range of respiratory pathogens can better inform public health preparedness for potential outbreaks and epidemics. Here, we simultaneously sequenced 38 pathogens including influenza viruses, coronaviruses and bocaviruses, to examine the abundance and seasonality of respiratory pathogens in urban wastewater. We deployed a targeted bait capture method and short-read sequencing (Illumina Respiratory Virus Oligos Panel; RVOP) on composite wastewater samples from 8 wastewater treatment plants (WWTPs) and one associated hospital site. By combining seasonal sampling with whole genome sequencing, we were able to concurrently detect and characterise a range of common respiratory pathogens, including SARS-CoV-2, adenovirus and parainfluenza virus. We demonstrated that 38 respiratory pathogens can be detected at low abundances year-round, that hospital pathogen diversity is higher in winter vs. summer sampling events, and that significantly more viruses are detected in raw influent compared to treated effluent samples. Finally, we compared detection sensitivity of RT-qPCR vs. next generation sequencing for SARS-CoV-2, enteroviruses, influenza A/B, and respiratory syncytial viruses. We conclude that both should be used in combination; RT-qPCR allowed accurate quantification, whilst genomic sequencing detected pathogens at lower abundance. We demonstrate the valuable role of wastewater genomic surveillance and its contribution to the field of wastewater-based epidemiology, gaining rapid understanding of the seasonal presence and persistence for common respiratory pathogens. By simultaneously monitoring seasonal trends and early warning signs of many viruses circulating in communities, public health agencies can implement targeted prevention and rapid response plans.

1. Introduction

The use of whole genome sequencing has proved invaluable to the understanding and management of the SARS-CoV-2 pandemic, enabling the identification of the early lineage of the RNA virus in Wuhan (Wu et al., 2020), and the lineages and variants that have since emerged (Brunner et al., 2022). The approach to sequence emerging variants in an attempt to get ahead of the pandemic has been adopted across the world (Brito et al., 2022; Priesemann et al., 2021). Identifying and tracking new outbreaks has facilitated government decision making (Colton et al., 2023) and aided to the success and timing of lockdown

measures across the UK (Hillary et al., 2021). Initial method development for variant sequencing has been centred on whole genome sequencing from clinical isolates, giving insight into both the evolution and the epidemiology of the pandemic (Harvey et al., 2021).

The shedding of SARS-CoV-2 in the faeces of infected (symptomatic and asymptomatic) individuals (Brumfield et al., 2022; Schmitz et al., 2021) allows for downstream detection in wastewater, potentially capturing viruses from an entire population on the sewage network (Schmitz et al., 2021), regardless of symptoms or access to healthcare facilities. Hence, wastewater-based epidemiology (WBE), i.e., the analysis of wastewater for viral quantities to understand increases and

* Corresponding author.

E-mail address: rachel.williams@bangor.ac.uk (R.C. Williams).

<https://doi.org/10.1016/j.watres.2024.121612>

Received 28 November 2023; Received in revised form 11 April 2024; Accepted 12 April 2024

Available online 13 April 2024

0043-1354/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

decreases of infections within the population, has been applied in many countries as a complementary surveillance tool to clinical testing (Monteiro et al., 2022; Prado et al., 2021). Applying sequencing methods to wastewater was instrumental in fast and efficient surveillance of communities (Li et al., 2022), due to the removal of testing bias and the detection of asymptomatic individuals. For example, an increase in SARS-CoV-2 infections was detected in wastewater 4–10 days earlier than those seen in clinical testing (Peccia et al., 2020; Wu et al., 2022), allowing time for appropriate risk management to be implemented, such as regional restrictions and the introduction of non-pharmaceutical interventions (NPIs e.g. social distancing). Thus, wastewater surveillance of SARS-CoV-2 has been shown to provide an effective early warning system for the emergence of new variants, or outbreaks in a specific geographical area.

As a result of wastewater monitoring, a global sampling and reporting infrastructure has been established during the COVID-19 pandemic. This has created a unique opportunity to leverage the experience gained in the pandemic and incorporate additional pathogens into surveillance programmes (Diamond et al., 2022). Respiratory pathogens create a large global burden year-round (GBD 2019; Diseases and Injuries Collaborators, 2020; Jin et al., 2021), and ‘influenza season’ is experienced in both hemispheres (Fisman, 2012). The co-occurrence of respiratory viruses such as coronaviruses, adenoviruses, rhinoviruses, RSV and influenza in infected patients has been documented (Ding et al., 2020; Mehta et al., 2021), and co-infections with different SARS-CoV-2 variants are thought to be one of the drivers of SARS-CoV-2 evolution (Rockett et al., 2022b). Increasing demand is therefore placed on understanding respiratory pathogen prevalence in a community, to enable efficient treatment administration and cost-effective management (Mehta et al., 2021).

Here, we explore the utility of targeted bait capture and short-read sequencing for WBE. To do this, we investigated respiratory pathogen prevalence in wastewater through comparative assessments (Fig. 1) with the following aims: 1, directly compare genomic sequencing and quantitative reverse transcription PCR (RT-qPCR) to evaluate their relative sensitivities and utilities for pathogen detection; 2, investigate seasonal

impact on pathogen presence with respect to the known seasonality of viral infections (Fisman, 2012; Moriyama et al., 2020); 3, compare methods for processing wastewater samples, and the impact of processing on whole genome pathogen recovery; 4, examine pathogen persistence through full-scale wastewater treatment plant (WWTP) processing, sequencing pathogens detected in raw sewage influent vs. treated effluent.

Overall, we show that wastewater genomic surveillance can be used to provide a non-biased approach to pathogen detection, over a large geographic area. The discoveries reported here substantially advance WBE – a field with increasing importance for public health.

2. Methods

2.1. Sample collection and processing

Samples were collected from eight urban wastewater treatment plants (WWTPs) serving towns and cities across North Wales and one large municipal hospital (Fig. 1). Samples were collected in the Summer (July/August) and Winter (November/December) of 2022 (Table S1). WWTP samples were untreated raw influent, and hospital samples were untreated raw near-source samples. Both were sampled via refrigerated autosamplers which collected 50 ml every 15 min to create 24 h composites. In addition, to screen for respiratory pathogens exiting WWTPs, a small number of influent and treated effluent samples were collected on the same day (Table S3). Wastewater was transported chilled to the laboratory for processing. All samples were spiked with *Pseudomonas* virus Phi6 as a processing control, and then WWTP samples, and winter hospital samples, were processed using polyethylene glycol (PEG) precipitation (Farkas et al., 2021). In summary, 200 ml of sample was centrifuged, and 150 ml of supernatant combined with PEG/NaCl for a final concentration of 10 % and 2 %, respectively, after pH adjustment to 7–7.5. Following a 16 h incubation at 4 °C, samples were centrifuged, and viral nucleic acids were preserved in the pellet. To align with other ongoing wastewater monitoring programs at the time, samples collected in the summer from the hospital were processed using ammonium

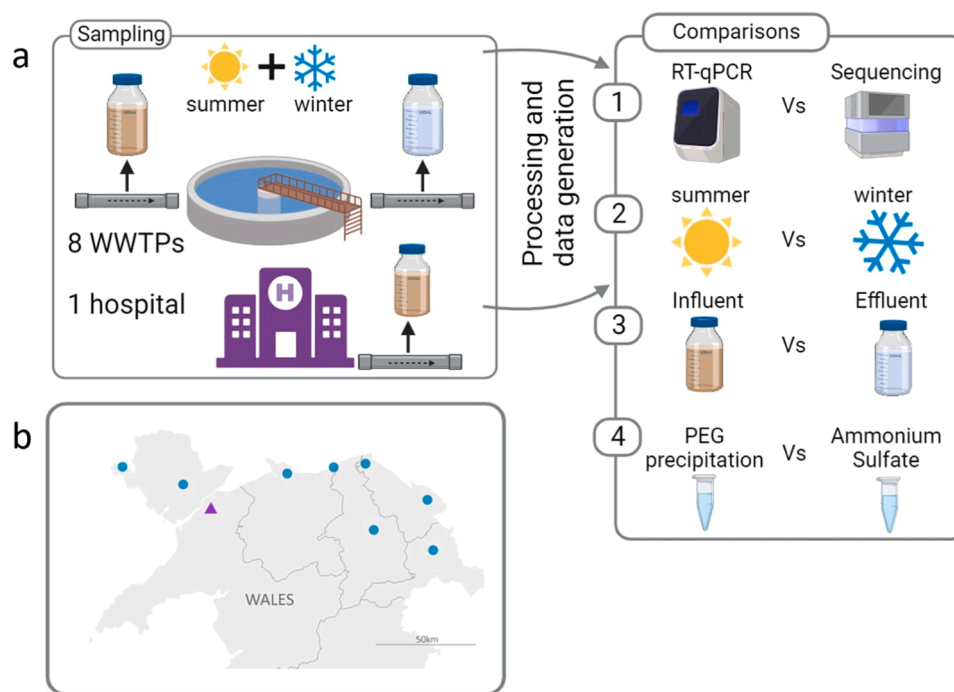


Fig. 1. Sample collection and processing. a: Overview of the sampling process and the subsequent comparisons made. Effluent samples were taken from one of the eight Wastewater Treatment Plants. Figure made using BioRender. b: The distribution of sampling sites across North Wales. Locations were sampled in summer and then re-sampled in winter. Samples were taken from Wastewater Treatment Plants (blue circles) and one hospital (purple triangle).

sulfate (AS) precipitation (Kevill et al., 2022). Briefly, 150 ml supernatant was combined with 57 g of ammonium sulfate (AS; Sigma-Aldrich, USA). Following AS dissolution, samples were incubated for 1 h at 4 °C, before 30 min centrifugation and pellet resuspension. Total nucleic extraction took place with 0.1 ml on a KingFisher 96 Flex system (Thermo Scientific, USA) using Nuclisens easyMAG extraction reagents (BioMerieux, France).

2.2. Quantification and library preparation

All samples were screened for SARS-CoV-2 using reverse transcription quantitative PCR (RT-qPCR), and these data were used to select samples for genomic sequencing. SARS-CoV-2 was quantified using RT-qPCR on a QuantStudio Flex 6 Real-Time PCR System (Applied Biosystems, USA) and the N1 primer probe set (CDC, 2020). Additional assays targeting *Enterovirus* spp., influenza A, influenza B viruses and respiratory syncytial virus (RSV; Farkas et al., 2022) were run on all samples from the winter collection. RT-qPCR reaction mixes contained TaqMan viral 1-step RT-qPCR master mix (Applied Biosystems Inc., USA), 1 µg bovine serum albumin (BSA), 10 µM forward, 20 µM reverse primers and 5 µM probe. Total RNA was quantified using a Qubit 4 Fluorometer (Thermo Scientific, USA).

Sequencing libraries were prepared using Illumina RNA Prep with Enrichment and IDT for Illumina DNA/RNA UD indexes. RNA denaturation, synthesis of first and second strand cDNA, cDNA tagmentation, cleaning, and normalising the libraries were all carried out following the manufacturer guidelines. Samples were enriched as single-plex reactions with the Respiratory Virus Oligos Panel (RVOP) and then pooled equimolarly. The hybridisation step was left overnight at 58 °C for maximum hybridisation. Hybridised probes were captured before amplification consisting of 98 °C for 30 s, 14 X (98 °C for 10 s, 60 °C for 30 s, 72 °C for 30 s), before a 5 min hold of 72 °C. Quantification and quality control of libraries was performed using a TapeStation (Agilent, USA) with D1000 reagents, and a Qubit BR DNA assay kit.

2.3. Sequencing and analysis

The library pool was spiked with 3 % PhiX, an adapter ligated library supplied as a sequencing quality control by Illumina. The final library was loaded at 650pM and run on a NextSeq1000 System using a P2 kit (2 × 150 bp) following Illumina concentration loading guidelines. FASTQ data files were used in Illumina's DRAGEN RNA Pathogen Detection Pipeline for analysis and viral detection. The pipeline uses reference genomes to analyse pathogen data and create consensus FASTAs, also outputting percentages of pathogen genomes with >5-times coverage. Pathogens without sufficient coverage to generate a consensus sequence were removed from the dataset. Resulting data were further analysed and visualised using R (R Core Team, 2022). Consensus sequences were generated within the DRAGEN RNA Pathogen Detection Pipeline for user specified pathogens. The BLASTN search tool in the NCBI database was used to compare consensus sequence similarity with the NCBI archives (Altschul et al., 1990).

2.4. Method comparison

Wastewater samples collected from the hospital in the summer were processed by ammonium sulfate (AS) precipitation to align with other national wastewater monitoring programs. Winter samples were processed using polyethylene glycol (PEG) precipitation to align with updated methods of surveillance. To ensure that comparisons could be made between these datasets due to different processing methods, six winter samples were processed in duplicate, once with PEG and once with AS. They were sequenced in the winter run to eliminate differences in performance between sequencing runs. A chi-square test of independence was performed to examine the relationship between the processing method, and the number of pathogens detected. Mann-Whitney

U tests were used to determine the statistical significance in differences between comparisons (Fig. 1).

3. Results

3.1. Sample selection and quality check

We aimed to determine the presence and seasonality of common respiratory pathogens in wastewater across North Wales. We sampled eight wastewater treatment plants (WWTPs) and one hospital site in the summer and the following winter. Post processing, RNA recovery was calculated at ~ 5 %, as determined from the recovery of *Pseudomonas* virus Phi6 spiked into the sample before processing.

The change in the processing of hospital samples between sampling periods was used as an opportunity to test the effect of viral precipitation methods on sequencing success by processing duplicate samples. Despite a larger number of pathogens being detected in PEG vs. AS precipitated samples (46 vs. 31, respectively), no significant difference was found between the processing method and the number of pathogens detected, or the coverage of the pathogen sequenced ($P > 0.05$).

3.2. RT-qPCR and sequencing comparison

We compared the RT-qPCR and genomic sequencing approaches with respect to their sensitivity for virus detection (Fig. 2). SARS-CoV-2 was detected in all samples tested with RT-qPCR, with virus concentration ranging from 10^2 to 10^5 gc/µl. However, only 63 % of samples successfully detected SARS-CoV-2 by sequencing, i.e., 37 % negative results.

We also used RT-qPCR to assess the seasonal impact on the concentration of SARS-CoV-2. The analysis revealed a significant seasonal effect ($P < 0.05$), with median concentrations of SARS-CoV-2 at 1.1×10^4 gc/L in winter compared to 3.3×10^2 gc/L in summer (Fig. 3). In contrast, sequencing did not demonstrate any significant seasonal differences.

In assays to detect *Enterovirus* spp., influenza A and B viruses and RSV, we observed greater sensitivity with sequencing than with RT-qPCR (Fig. 2). Enteroviruses were detected in all samples tested using sequencing, but only one sample was shown as positive using RT-qPCR. Influenza A virus was detected in 75 % of samples using both sequencing and RT-qPCR, however only 50 % of samples overlapped in successful detection in both methods. Influenza B virus was not detected by RT-qPCR but was successfully detected in 63 % of samples using sequencing. RSV was more difficult to detect by either method: all

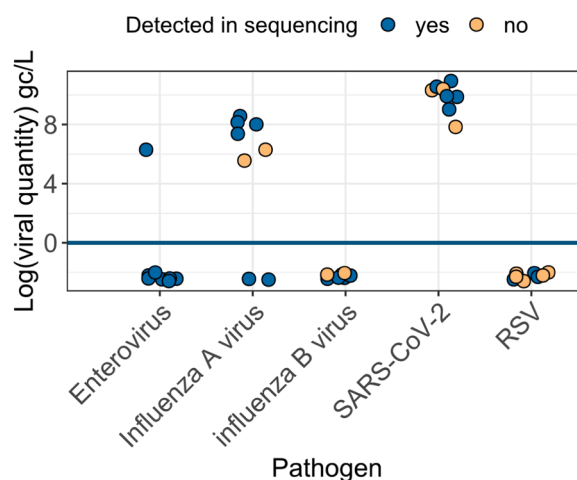


Fig. 2. RT-qPCR quantities for a subset of viruses in winter wastewater samples. Points above the line represent log(gc/L), points below the line are negative for RT-qPCR. Colour represents success of detection in sequencing.

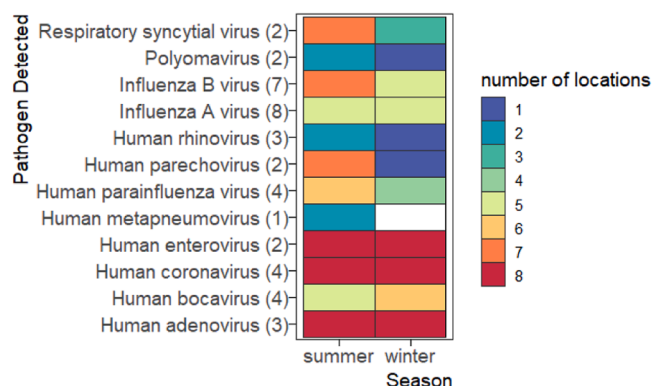


Fig. 3. Pathogens present across the number of wastewater treatment plants (maximum of 8) identified through sequencing. Summarised pathogens are shown, the total number of strains included are shown in brackets (included in full in the Supplementary Information).

samples were negative as tested by RT-qPCR, but 25 % were positive by sequencing.

3.3.1. Seasonality of respiratory pathogens using short read sequencing

Evaluating the presence of respiratory pathogens in communities across seasons can better prepare public health services for a rapid response to future outbreaks or epidemics. Sampling wastewater from the same locations in summer and winter allowed us to assess the seasonal presence of a pathogen detected at WWTPs across North Wales. Across sites, we observed a larger number of respiratory viruses in summer vs. winter (132 vs. 106, respectively) in the sequencing data (Fig. 3); however, there was no significant difference in the diversity of viruses (i.e. taxonomic groups of viruses and/or strains within a virus) present in summer vs. winter or the genomic coverage of the pathogens detected ($P > 0.05$). Human adenovirus was detected at all eight locations both in summer and winter. However, we saw seasonal variation in influenza B virus and human parechovirus, where the latter was detected at seven sites in summer and only at one site in winter.

3.3.2. Hospital wastewater monitoring

Hospital samples collected in the summer were processed using AS precipitation to align with other national wastewater monitoring programs, whereas winter samples were processed using PEG precipitation to align with updated methods of surveillance. No significant differences in either the number of pathogens detected ($P > 0.05$) or the genome coverage of pathogens ($P > 0.05$) was found between the two methods (AS vs. PEG) comparing summer and winter sampling events at the

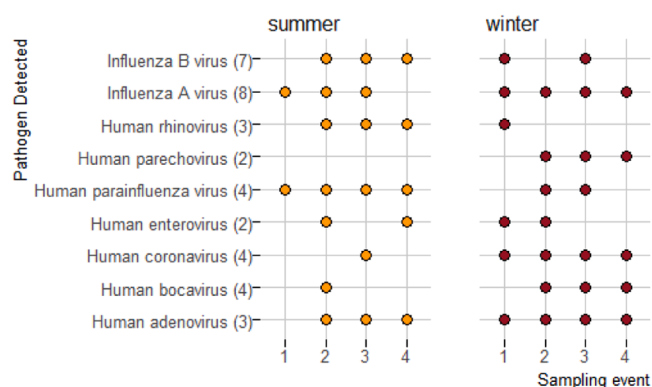


Fig. 4. Pathogens found in the hospital wastewater in summer and winter across four sampling events/days each. Summarised pathogens are shown, the total number of strains included are shown in brackets (included in full in the Supplementary Material).

hospital (Fig. 4).

A higher number of pathogens was detected in winter in hospital wastewater, but no significant difference was found between seasons ($P > 0.05$), corroborating results found at WWTPs (Fig. 3). There were differences in patterns observed at WWTPs and at the hospital; for example, human parechovirus was not detected in the summer but was captured and detected in 75 % of the sampling events at the hospital in the winter (Fig. 4). This pattern was not seen at the WWTPs, which showed a higher abundance of the pathogen in the summer. Unsurprisingly, adenoviruses, coronaviruses and influenza A virus were detected in all sampling events through the winter.

3.4. Detection of pathogens exiting WWTPs

We sequenced raw influent and treated effluent wastewater from one WWTP to assess the likelihood of specific pathogens being detected after wastewater treatment plant processing (Fig. 5). Influent and effluent composite samples were collected on the same day, on three separate days, creating three sampling events of both influent and effluent. Samples were selected based on a positive effluent result for SARS-CoV-2 using RT-qPCR, and then paired with their corresponding influent sample (all influent samples collected from this site were positive for SARS-CoV-2 around this time). The RVOP kit performed poorly in its ability to sequence SARS-CoV-2 in either the influent or the effluent samples. Overall, we detected significantly higher virus levels in influent compared to effluent samples ($P < 0.05$). This indicates that sewage treatment plant processes are successfully eliminating/substantially reducing the concentration of a range of pathogens, including human metapneumovirus and strains of human bocavirus and influenza viruses. Conversely, the human adenoviruses were detected and sequenced in both influent and effluent samples, although these were not enteric adenoviruse strains (Table S3).

4. Discussion

In this study we assessed the usefulness and applicability of short-read sequencing of human viruses for wastewater monitoring and surveillance. Overall, the RVOP kit was robust and successful for the preliminary assessment of viral abundance, seasonality and diversity in wastewater from different sources. We compared tools that are valuable to the field of WBE, to assess their contribution to monitoring a wide range of pathogens. While sequencing was able to detect viruses missed

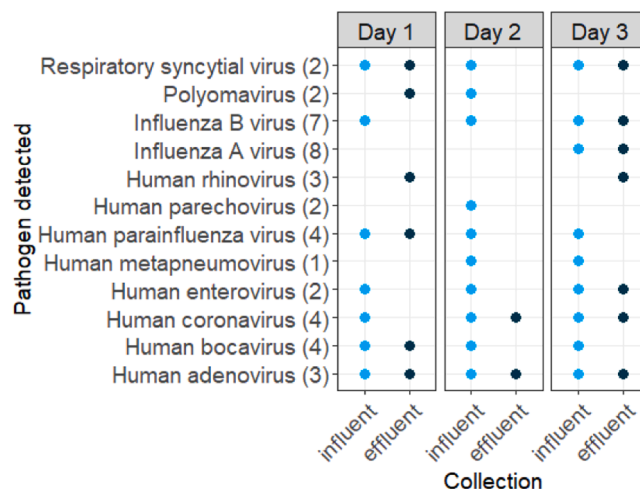


Fig. 5. Comparison of pathogens detected before and after wastewater treatment plant processing at three sampling events. Influent and effluent samples were 24-hour composites. Summarised pathogens are shown, the total number of strains included are shown in brackets (see Supplementary Material for detail).

by RT-qPCR, like Influenza B virus, RT-qPCR demonstrated greater sensitivity for SARS-CoV-2, which is likely a reflection of assay sensitivity. There are differences in the assays' limits of detection (LOD) - 787 gc/ μ L and 611 gc/ μ L for Influenza B virus and SARS-CoV-2, respectively (Farkas et al., 2022), but it is likely that assay sensitivity rather than an inherent superiority of sequencing-based methods, seems to drive observed detection differences between the methods.

Carefully matching assay sensitivity to pathogen levels under surveillance is important in optimizing RT-qPCR vs. short-read sequencing approaches. Sample concentration also impacts detection; samples with high target virus concentrations will typically yield more sequencing data than low concentration samples. In this study, each sample was individually balanced before pooling to improve coverage equally across all samples. This approach during sequencing library preparation allowed us to infer presence/absence of pathogens, but not their concentrations. Samples that were successfully detected in RT-qPCR could be quantified, due to the inclusion of standards that were of known concentration, included in each qPCR run. We demonstrate here that sequencing was able to detect pathogens at lower abundance, while RT-qPCR was complementary allowing accurate quantification. Quantitative data provides estimates of population level infection, whereas sequencing data provides a complementary type of evaluation, identifying new and existing mutations in circulation within an infected population (Khan et al., 2023; Li et al., 2022). From the perspective of community surveillance, both are important factors to consider in management of public health, and both should be used in combination.

4.1. Genomic surveillance of seasonal variation

Seasonal variation in respiratory pathogens is already well documented (Fisman, 2012; Lofgren et al., 2007), with general conclusions confirming seasonal patterns for some species, but numerous areas of uncertainty exist. For example, transmission of influenza viruses is typically higher in the winter season due to weakened immune system of the host, poor ventilation and more time spent indoors in crowded environments (Lofgren et al., 2007). Here, we noted high detection rates of influenza viruses using sequencing regardless of the season (Figs. 2-3), which may be due to international travel. If winter travel were responsible for the respiratory pathogen peak seen each winter (Cherrie et al., 2018), we would expect a larger summer peak correlating with a typical 'summer holiday'. The possibility of a host-based defence mechanism during the summer has been considered (Dowell, 2001), with the effect of environmental factors on the host airway defence mechanisms a contributing factor to low case rates in the summer (Moriyama et al., 2020).

We found limited seasonal patterns in the abundance of enteroviruses, adenoviruses and polyomaviruses, as noted previously in clinical studies (Liu et al., 2016; Okada et al., 2013; Pons-Salort et al., 2015) and in wastewater (Brinkman et al., 2017; Farkas et al., 2020). Coronaviruses and bocaviruses were more abundant in wastewater collected during the winter as opposed to the summer (Foulongne et al., 2006; Nichols et al., 2021); however, it has been suggested that bocavirus lacks seasonal patterns (Bastien et al., 2006; Maggi et al., 2007). In contrast, RSV, parechovirus and parainfluenza virus were more common in wastewater during the summer than winter, corroborating previous clinical findings (Lee et al., 2023; Li et al., 2019), but contradicting findings of a higher abundance generally for respiratory viruses in winter reported by Cherrie et al. (2018). This may be due to international travel, as discussed above, the varied distribution of viral strains, or the fact that clinical studies are based on data derived from hospitalised patients, excluding infected individuals with mild symptoms (Obando-Pacheco et al., 2018; Tapia et al., 2008). Further, the 2022 to 2023 winter season saw a decrease in total respiratory infection incidents largely due to fewer reported COVID-19 cases (UK Health Security Agency, 2023). Influenza and other respiratory virus incidents increased compared to the previous season, although remained low

relative to pre-pandemic years. The prior 2021–2022 season had seen suppressed circulation of seasonal viruses due to ongoing pandemic control measures (Hillary et al., 2021).

4.2. Detection for WBE application

The year-round persistence of SARS-Cov-2 has been documented (Khan et al., 2023), demonstrating the strength of the virus to persist and infect potential hosts. Data presented here, and data from sequencing panels in general, can be used to provide information about viruses circulating in a community. As a result of the acknowledged limitations of wastewater surveillance, such as spatial resolution or sensitivity (Diamond et al., 2022; Keshaviah et al., 2023), WBE has typically taken a complementary role to more conventional methods of public health surveillance (O'Keefe, 2021). Technological advances, however, now allow us to examine a broad range of pathogens using next generation sequencing, providing us with the opportunity to use the data to guide public action, rather than just complement it (Khan et al., 2023; Ogunbayo et al., 2023). For example, screening a community for both transient and permanent infections using whole genome coverage can provide a baseline for pathogens circulating, it can facilitate epidemiologic investigations into vaccine effectiveness or even inform vaccination strategies (Graf et al., 2016). It may also provide insights into the asymptomatic carriage rate of respiratory pathogens.

While the insights gained from whole genome sequencing are greater than ever before, there are several barriers to rolling out these methods for routine diagnostics such as the costs and the complexity of data analysis (Kapel et al., 2023). Methods are typically developed for, and used on, clinical samples (Li et al., 2019; Ogunbayo et al., 2023; Rockett et al., 2022b). The wastewater matrix adds further complexity for biological assays; for example due to limitors such as inhibitors and active enzymatic processes. Nevertheless, the methods are becoming increasingly refined and harmonised to the extent that wastewater testing and WBE are now more established as a standard practice (Brunner et al., 2022; Tisza et al., 2023).

Here, we have demonstrated that pathogens are detectable in wastewater, even outside their typical season of prevalence, and that whole genome sequencing can allow for fine scale detection of a wide range of respiratory pathogens in wastewater. The variety of pathogens detected demonstrates the effectiveness of wastewater surveillance in the monitoring of respiratory pathogens. The transport times of wastewater through wastewater treatment plants typically range from several hours to a day or more, depending on the efficiency and complexity of the treatment processes involved (Rieckermann et al., 2005; Sokáč and Velísková, 2016; Sonnenwald et al., 2023) The incongruence between WWTP and hospital data sets could be a result of the former including virus shed by individuals with subclinical infections. Furthermore, differences in handling of waste at the source can be a contributing factor. For example, hospital cleaning routines are likely to be more stringent than those of the average household, and chemicals (e.g. disinfectants and detergents) added to water would likely have an impact on pathogens detected through sequencing (Ke et al., 2022). Leveraging the distinct profiles of pathogens detected across hospital, influent, and effluent sites through genomic surveillance provides an opportunity to track subclinical infections in communities and improve wastewater treatment procedures, for example to reduce pathogen levels in WWTP effluents.

5. Conclusions

Utilising a multi-pathogen detection panel, we demonstrate successful sequencing of 38 respiratory viruses from wastewater, enabling broad community surveillance capabilities. The application of targeted bait capture and short-read sequencing provides resolution down to the lineage level for pathogens of concern, while concurrently monitoring diverse threats for public health safety. This work validates wastewater-

based sequencing as a powerful public health tool, providing a near real-time snapshot of pathogens circulating in communities. By integrating sequencing with RT-qPCR, we demonstrate a comprehensive surveillance workflow to detect emerging outbreaks and advance WBE.

CRedit authorship contribution statement

Rachel C. Williams: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Kata Farkas:** Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing. **Alvaro Garcia-Delgado:** Data curation, Writing – review & editing. **Latifah Adwan:** Data curation. **Jessica L. Kevill:** Conceptualization, Writing – review & editing. **Gareth Cross:** Funding acquisition. **Andrew J. Weightman:** Funding acquisition, Writing – review & editing. **Davey L. Jones:** Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This research was supported by Welsh Government under the Welsh Wastewater Programme (C035/2021/2022). We thank Tony Harrington and Ian Trick at Dŵr Cymru Welsh Water, UK for their assistance in wastewater sample collection and Bangor laboratory team for their assistance with sample preparation. We thank the Welsh Wastewater Monitoring Group, Bangor and Cardiff, UK for their assistance in the development of sample processing. We thank Steve Copley at Welsh Government for project supervision and management.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.watres.2024.121612](https://doi.org/10.1016/j.watres.2024.121612).

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Bastien, N., Brandt, K., Dust, K., Ward, D., Li, Y., 2006. Human bocavirus infection. *Canada. Emerg Infect Dis.* 12, 848–850. <https://doi.org/10.3201/eid1205.051424>.
- Brinkman, N.E., Fout, G.S., Keely, S.P., 2017. Retrospective surveillance of wastewater to examine seasonal dynamics of enterovirus infections. *mSphere* 2, e00099. <https://doi.org/10.1128/msphere.00099-17>.
- Brito, A.F., Semenova, E., Dudas, G., Hassler, G.W., Kalinich, C.C., Kraemer, M.U.G., Ho, J., Tegally, H., Githinji, G., Agoti, C.N., Matkin, L.E., Whittaker, C., Howden, B. P., Sintchenko, V., Zuckerman, N.S., Mor, O., Blankenship, H.M., de Oliveira, T., Lin, R.T.P., Siqueira, M.M., Resende, P.C., Vasconcelos, A.T.R., Spilki, F.R., Aguiar, R.S., Alexiev, I., Ivanov, I.N., Philipova, I., Carrington, C.V.F., Sahadeo, N.S. D., Branda, B., Gurry, C., Maurer-Stroh, S., Naidoo, D., von Eije, K.J., Perkins, M.D., van Kerkhove, M., Hill, S.C., Sabino, E.C., Pybus, O.G., Dye, C., Bhatt, S., Flaxman, S., Suchard, M.A., Grubaugh, N.D., Baele, G., Faria, N.R., 2022. Global disparities in SARS-CoV-2 genomic surveillance. *Nat. Commun.* 13, 7003. <https://doi.org/10.1038/s41467-022-33713-y>.
- Brumfield, K.D., Leddy, M., Usmani, M., Cotruvo, J.A., Tien, C.T., Dorsey, S., Graubics, K., Fanelli, B., Zhou, I., Registe, N., Dadlani, M., Wimalarante, M., Jinasena, D., Abayagunawardena, R., Withanachchi, C., Huq, A., Jutla, A., Colwell, R.R., 2022. Microbiome analysis for wastewater surveillance during COVID-19. *mBiomBio* 13. <https://doi.org/10.1128/mbio.00591-22>.
- Brunner, F.S., Brown, M.R., Bassano, I., Denise, H., Khalifa, M.S., Wade, M.J., van Aerle, R., Kevill, J.L., Jones, D.L., Farkas, K., Jeffries, A.R., Cairns, E., Wierzbicki, C., Paterson, S., 2022. City-wide wastewater genomic surveillance through the successive emergence of SARS-CoV-2 Alpha and Delta variants. *Water. Res.* 226, 1–9. <https://doi.org/10.1016/j.watres.2022.119306>.
- Cherrie, M.P.C., Nichols, G., Iacono, G.Lo, Sarra, C., Hajat, S., Fleming, L.E., 2018. Pathogen seasonality and links with weather in England and Wales: a big data time series analysis. *BMC Public Health* 18. <https://doi.org/10.1186/s12889-018-5931-6>.
- Colton, H., Parker, M.D., Stirrup, O., Blackstone, J., Loose, M., McClure, C.P., Roy, S., Williams, C., McLeod, J., Smith, D., Taha, Y., Zhang, P., Hsu, S.N., Kele, B., Harris, K., Mapp, F., Williams, R., Flowers, P., Breuer, J., Partridge, D.G., de Silva, T. I., 2023. Factors affecting turnaround time of SARS-CoV-2 sequencing for inpatient infection prevention and control decision making: analysis of data from the COG-UK HOCl study. *J. Hospital Infect.* 131, 34–42. <https://doi.org/10.1016/j.jhin.2022.09.022>.
- Diamond, M.B., Keshaviah, A., Bento, A.I., Conroy-Ben, O., Driver, E.M., Ensor, K.B., Halden, R.U., Hopkins, L.P., Kuhn, K.G., Moe, C.L., Rouchka, E.C., Smith, T., Stevenson, B.S., Susswein, Z., Vogel, J.R., Wolfe, M.K., Stadler, L.B., Scarpino, S.V., 2022. Wastewater surveillance of pathogens can inform public health responses. *Nat. Med.* 28, 1992–1995. <https://doi.org/10.1038/s41591-022-01940-x>.
- Ding, Q., Lu, P., Fan, Y., Xia, Y., Liu, M., 2020. The clinical characteristics of pneumonia patients coinfecting with 2019 novel coronavirus and influenza virus in Wuhan. *China. J. Med. Virol.* 92, 1549–1555. <https://doi.org/10.1002/jmv.25781>.
- Dowell, S.F., 2001. Seasonal variation in host susceptibility and cycles of certain infectious diseases. *Emerg. Infect. Dis.* 7, 369–374. <https://doi.org/10.3201/eid0703.010301>.
- Farkas, K., Hillary, L.S., Malham, S.K., McDonald, J.E., Jones, D.L., 2020. Wastewater and public health: the potential of wastewater surveillance for monitoring COVID-19. *Curr. Opin. Environ. Sci. Health.* <https://doi.org/10.1016/j.coesh.2020.06.001>.
- Farkas, K., Hillary, L.S., Thorpe, J., Walker, D.L., Lowther, J.A., McDonald, J.E., Malham, S.K., Jones, D.L., 2021. Concentration and quantification of sars-cov-2 RNA in wastewater using polyethylene glycol-based concentration and qRT-PCR. *Methods Protoc.* 4, 1–9. <https://doi.org/10.3390/MP54010017>.
- Farkas, K., Pellett, C., Alex-Sanders, N., Bridgman, M.T.P., Corbishley, A., Grimsley, J.M. S., Kasprzyk-Hordern, B., Kevill, J.L., Pantea, I., Richardson-O'Neill, I.S., Lambert-Slosarska, K., Woodhall, N., Jones, D.L., 2022. Comparative assessment of filtration and precipitation-based methods for the concentration of SARS-CoV-2 and other viruses from wastewater. *Microbiol. Spectr.* 10, e01102–e01122. <https://doi.org/10.1128/spectrum.01102-22>.
- Fisman, D., 2012. Seasonality of viral infections: mechanisms and unknowns. *Clin. Microbiol. Infect.* 18, 946–954. <https://doi.org/10.1111/j.1469-0691.2012.03968.x>.
- Foulongne, V., Olejnik, Y., Perez, V., Elaerts, S., Rodière, M., Segondy, M., 2006. Human bocavirus in French children. *Emerging Infect. Dis.* 12, 1251–1253.
- GBD 2019 Diseases and Injuries Collaborators, 2020. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 396, 1204–1222. [https://doi.org/10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9).
- Graf, E.H., Simmon, K.E., Tardif, K.D., Hymas, W., Flygare, S., Eilbeck, K., Yandell, M., Schlager, R., 2016. Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: a systematic comparison to a commercial PCR panel. *J. Clin. Microbiol.* 54, 1000–1007. <https://doi.org/10.1128/JCM.03060-15>.
- Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S.J., Robertson, D.L., 2021. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* 19, 409–424. <https://doi.org/10.1038/s41579-021-00573-0>.
- Hillary, L.S., Farkas, K., Maher, K.H., Lucaci, A., Thorpe, J., Distaso, M.A., Gaze, W.H., Paterson, S., Burke, T., Connor, T.R., McDonald, J.E., Malham, S.K., Jones, D.L., 2021. Monitoring SARS-CoV-2 in municipal wastewater to evaluate the success of lockdown measures for controlling COVID-19 in the UK. *Water. Res.* 200. <https://doi.org/10.1016/j.watres.2021.117214>.
- Jin, X., Ren, J., Li, R., Gao, Y., Zhang, H., Li, J., Zhang, J., Wang, X., Wang, G., 2021. Global burden of upper respiratory infections in 204 countries and territories, from 1990 to 2019. *EclinicalMedicine* 37. <https://doi.org/10.1016/j.eclinm.2021.100986>.
- Kapel, N., Kalimeris, E., Lumley, S., Decano, A., Rodger, G., Lopes Alves, M., Dingle, K., Oakley, S., Barrett, L., Barnett, S., Crook, D., Eyre, D.W., Matthews, P.C., Street, T., Stoesser, N., 2023. Evaluation of sequence hybridization for respiratory viruses using the Twist Bioscience Respiratory Virus Research panel and the OneCodex Respiratory Virus sequence analysis workflow. *Microb. Genom.* 9. <https://doi.org/10.1099/mgen.0.001103>.
- Ke, M., Xu, N., Zhang, Z., Qiu, D., Kang, J., Lu, T., Wang, T., Peijnenburg, W.J.G.M., Sun, L., Hu, B., Qian, H., 2022. Development of a machine-learning model to identify the impacts of pesticides characteristics on soil microbial communities from high-throughput sequencing data. *Environ. Microbiol.* 24, 5561–5573. <https://doi.org/10.1111/1462-2920.16175>.
- Keshaviah, A., Diamond, Megan B., Wade, M.J., Scarpino, S.V., Ahmed, W., Amman, F., Aruna, O., Badilla-Aguilar, A., Bar-Or, I., Berghaler, A., Bines, J.E., Bivins, A.W., Boehm, A.B., Brault, J.M., Burnett, J.B., Chapman, J.R., Chaudhuri, A., de Rodo Husman, A.M., Delatolla, R., Dennehy, J.J., Diamond, Megan Beth, Donato, C., Duizer, E., Egwuenu, A., Erster, O., Fatta-Kassinos, D., Gaggero, A., Gilpin, D.F., Gilpin, B.J., Graber, T.E., Green, C.A., Handley, A., Hewitt, J., Holm, R.H., Insam, H., Johnson, M.C., Johnson, R., Jones, D.L., Julian, T.R., Jyothi, A., Kohn, T., Kuhn, K. G., La Rosa, G., Lesenfans, M., Manuel, D.G., D'Aoust, P.M., Markt, R., McGrath, J. W., Medema, G., Moe, C.L., Murni, I.K., Naser, H., Naughton, C.C., Ogorzaly, L., Oktaria, V., Ort, C., Karaolia, P., Patel, E.H., Paterson, S., Rahman, M., Rivera-Navarro, P., Robinson, A., Santa-Maria, M.C., Schmitt, H., Smith, T., Stadler, L.B., Stassijns, J., Stenico, A., Street, R.A., Suffredini, E., Susswein, Z., Trujillo, M.,

