

Environmental Genomic Surveillance

Insights and Applications from SARS-CoV-2 Wastewater Analysis



Wastewater Analysis Group: Tyler Chafin^{1,2}, Thomas Cornulier¹, Zhou Fang¹, Anastasia Frantsuzova¹, Claus-Dieter Mayer¹, Sonia Mitchell¹, Jackie Potts¹
 1. Biomathematics and Statistics Scotland
 2. Corresponding author (Contact: tylerchafin@bioss.ac.uk)

Collaborators



Introduction

COVID-19 infection levels may be monitored via wastewater (WW) sampling of SARS-CoV-2 concentrations. The current sampling scheme in Scotland covers over 75% of the population. BioSS provides weekly reports to the Scottish Government, leveraging environmental and case data to objectively quantify COVID-19 prevalence. This methodological approach, while currently focused on COVID-19, showcases the potential of environmental surveillance (ES) for tracking diverse pathogens and application areas (e.g., monitoring antimicrobial resistance), positioning it as a versatile tool for science-informed policy across various sectors.

Genomic Surveillance

Genomic sequencing enhances environmental surveillance by uncovering the diversity and prevalence of viral lineages within the population. This offers robust spatial and temporal tracking of virus variants, and early detection of emerging strains. This approach captures underlying variation that shapes observed prevalence, thus informing targeted public health policy decisions such as outbreak response and risk assessment. Harnessing this data necessitates overcoming novel technical challenges, underscoring the need for continuous innovation.

Methods



1. Scottish Water (SW) samples water for testing (since Aug. 2020)
2. SEPA conducts RT-qPCR to quantify SARS-CoV-2 prevalence
3. NHS Lothian performs whole-genome sequencing (since Oct. 2022)
4. BioSS models prevalence (qPCR) and viral lineage abundance (genomics)

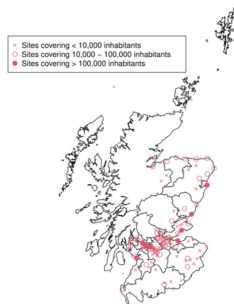


Figure 1 – Distribution of WW Sites in Scotland (left):

>100 sites are sampled for prevalence, and >90 for sequencing weekly, covering ~80% of the Scottish population

Results

Virus prevalence (from qPCR) and relative lineage abundance (from sequencing) provide critical insights into the dynamics of the pandemic. These data-driven tools allow us to track the temporal and spatial evolution of the virus and its variants, thereby informing strategic public health decisions, including timely interventions and resource allocation.

Figure 2 – SARS-CoV-2 Prevalence from WW:

This visualisation displays the temporal progression of virus prevalence within the population, calculated as a national average, weighted by population size and smoothed over 7-day periods. This representation measures the 'waves' of the pandemic, correlating with the rise and subsequent decline in case numbers.

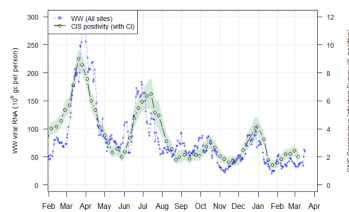
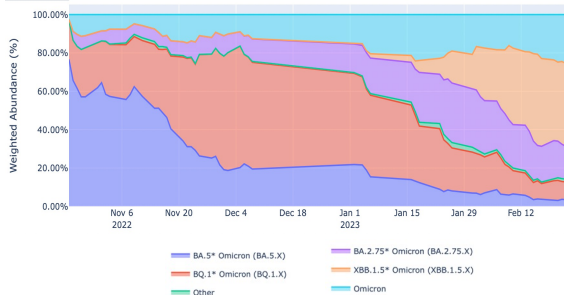


Figure 3 – SARS-CoV-2 Relative Lineage Abundance from WW:

Relative lineage prevalence through time reported as national average, weighted by population, and smoothed in 7-day windows, illustrating variant growth and turnover.



Research Directions

Enhancing Applicability through Methodological Innovation and Data Integration

- A primary avenue involves leveraging genomic variant data to refine prevalence models derived from RT-qPCR.
- We also see potential in enhancing prevalence forecasting by incorporating variants as exogenous features, thereby improving the predictive capacity of our models.
- A key challenge is to extend the methods to other pathogens, enabling a more versatile, pathogen-agnostic approach to environmental surveillance.
- Developing better approaches for early detection and for informing deployment of more costly surveillance methods, e.g., by pinpointing when and where they are most needed.
- Additionally, monitoring changes in PCR-based test specificity via wastewater variant tracking could provide valuable insights into potential shifts in testing efficacy

Conclusion

- Environmental surveillance presents challenges when applied in a pathogen-agnostic manner; while it may not offer a one-size-fits-all solution for "disease X", it excels when targeted towards specific pathogens of interest (....and can be deployed quickly + cost-effectively)
- The addition of genomic sequencing enhances the utility of environmental surveillance by allowing insights into viral lineage prevalence and dynamics
- Potential applications extend beyond public health into sectors such as agriculture (e.g., monitoring of agricultural slurries, soil), especially as technological advancements continue to drive down costs



Figure 4 – Example Outputs of Reporting Pipeline: Automated, interactive reporting enables streamlined visualisation and rapid interpretation by stakeholders while maintaining interpretability in high-level overviews



sefari.scot



info@sefari.scot



@SEFARIScot

RESAS

Rural & Environmental Science and Analytical Services



Scottish Government
Riaghaltas na h-Alba
gov.scot