UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Posgrados

Analysis of SARS-CoV-2 variants in wastewater from the Metropolitan District of Quito using a passive sampling 3D printed device

Tesis en torno a una hipótesis o problema de investigación y su contrastación

Paula Estefania Mejia Calle

Paúl Cárdenas A., M.D., Ph.D.
Director de Trabajo de Titulación

Trabajo de titulación de posgrado presentado como requisito para la obtención del título de Magister en Microbiología

Quito, 18 de noviembre del 2024

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

COLEGIO DE POSGRADOS

HOJA DE APROBACIÓN DE TRABAJO DE TITULACIÓN

Analysis of SARS-CoV-2 variants in wastewater from the Metropolitan District of Quito using a passive sampling 3D printed device

Paula Estefania Mejia Calle

Nombre del Director del Programa:	Patricio Rojas Silva
Título académico:	M.D., Ph.D.
Director del programa de:	Maestría en Microbiología

Nombre del Decano del colegio Académico:	Carlos Valle
Título académico:	Ph.D
Decano del Colegio:	COCIBA

Nombre del Decano del Colegio de Posgrados:	Darío Niebieskikwiat
Título académico:	Ph.D

Quito, noviembre del 2024

© DERECHOS DE AUTOR

Nombre del estudiante:	Paula Estefania Mejia Calle
Código de estudiante:	00331348
C.I.:	0150342517
Lugar y fecha:	Quito, 18 de noviembre de 2024

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en http://bit.ly/COPETheses.

UNPUBLISHED DOCUMENT

Note: The following graduation project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on http://bit.ly/COPETheses.

DEDICATORIA

A mi amada hija, Keyla Fierro,

Este logro es tanto tuyo como mío. Desde el primer momento, has sido mi mayor inspiración y motor para seguir adelante, incluso en los momentos más difíciles. Tu paciencia, esfuerzo y dedicación me dieron la fuerza para no rendirme, y aunque la distancia nos separó en ocasiones, siempre sentí tu amor y apoyo.

Superamos juntas los desafíos que la vida nos puso, y es por eso que esta tesis te pertenece tanto como a mí. Cada página está impregnada del cariño y sacrificio que compartimos. Gracias por ser mi luz en este camino y por enseñarme que, con amor y determinación, todo es posible.

Con todo mi amor,

Paula

AGRADECIMIENTOS

En primer lugar, agradezco a Dios por guiarme en este camino y permitirme encontrar esta maestría en el momento perfecto, su fortaleza me sostuvo en cada paso, dándome el valor para no rendirme ante las adversidades.

Agradezco a la Universidad San Francisco de Quito y al Instituto de Microbiología por brindarme un entorno enriquecedor para mi crecimiento profesional, así como por otorgarme una beca, sin ese apoyo, alcanzar esta meta no habría sido posible.

Expreso mi más profundo agradecimiento a mi tutor, Paúl Cárdenas Aldaz, por ser un mentor excepcional. Su apoyo constante y confianza en mi trabajo fueron esenciales para este proyecto, más que un tutor siempre estuvo dispuesto a ayudarme a superar las adversidades, confiando plenamente en mis habilidades.

A los miembros de mi comité, Verónica Barragán, Patricio Rojas y Manuel Baldeón, les agradezco por su compromiso, retroalimentación valiosa y por estar siempre pendientes del avance de este proyecto.

Al equipo de Bioinformática, especialmente a Mateo Carvajal, Erika Muñoz, Rommel Guevara y Denisse Benítez, gracias por su colaboración incondicional, por ayudarme a estandarizar los protocolos de laboratorio y pipelines bioinformáticos. Su paciencia y disposición fueron clave para el éxito de este trabajo.

A mis amigos de la cohorte 2022, gracias por ser un grupo tan unido y por hacer de esta experiencia algo tan especial e inolvidable.

Finalmente, quiero agradecer a mis padres, José Mejía y María Eulalia Calle, quienes siempre han estado a mi lado, apoyándome en cada paso de mi vida. Sin su amor y constante apoyo, este logro no habría sido posible.

RESUMEN

Introducción: La infección por SARS-CoV-2 en humanos produce la eliminación del ARN viral principalmente a través de gotitas respiratorias, pero también de forma activa a través de la materia fecal. La vigilancia de patógenos en aguas residuales se ha convertido en una herramienta esencial para la preparación ante epidemias, ha permitido el seguimiento de las variantes emergentes del SARS-CoV-2 en poblaciones con bajas tasas de diagnóstico clínico. Por esta razón, se implementó un enfoque de muestreo pasivo, que consiste en colocar dispositivos directamente en el flujo del agua durante un periodo prolongado, para obtener una muestra representativa del comportamiento del virus a lo largo del tiempo, sin la necesidad de una intervención activa o recolección frecuente a lo largo del día. Este enfoque contribuyo para determinar las variantes del SARS-CoV-2 en Quito, Ecuador, durante 2023-2024, para luego correlacionarlo con los resultados de la vigilancia genómica nacional realizada en pacientes en hospitales de Quito que proporcionan datos importantes en la vigilancia epidemiológica en Ecuador. **Métodos**: Se utilizó un equipo de muestreo pasivo tipo torpedo para recolectar las muestras, que contenía una membrana de nailon, gasa y un hisopo; el mismo que se colocó en colectores de aguas residuales de la PTAR Quitumbe, durante 24 horas semanales. Este método es un enfoque más sencillo para el muestreo de aguas residuales que el convencional que requiere varios litros de recolección. El ARN se concentró con PEG, se realizó extracción de ARN viral, seguida de PCR en tiempo real para detección de virus y PCR multiplex para secuenciación con Oxford Nanopore Technology y los datos bioinformáticos fueron analizados en Freyja. **Resultados**: Se observó un predominio de la variante ómicron, con linajes como HN.1, EG.5.1.6, XBB.1.5.1.5 y HV.1. En particular, se detectó el linaje JN.1 durante los picos de infección en Quito, en octubre de

2023 y febrero de 2024, en proporciones similares a las de las muestras clínicas circulantes en esos periodos, y en algunos casos, incluso antes de ser identificadas clínicamente. Cabe destacar que la metodología utilizada es altamente sensible a los cambios en los cebadores. Entre marzo y julio de 2024, los resultados no coincidieron con las muestras clínicas, ya que, aunque el SARS-CoV-2 seguía siendo detectable, solo se identificaron secuencias conservadas, lo que llevó a la asignación incorrecta de linajes, específicamente de linajes primitivos como JD.1. A partir de agosto de 2024, se comenzaron a detectar linajes recombinantes, siendo notable la aparición del linaje KP.2.3 en septiembre, correlacionándose con los hallazgos en muestras humanas. **Conclusión**: La vigilancia epidemiológica de las aguas residuales del SARS-CoV-2 es un enfoque sencillo para la caracterización genómica de las variantes circulantes y la cuantificación de la abundancia viral.

Palabras clave: Sars-CoV-2, equipo de muestreo pasivo tipo torpedo, vigilancia epidemiológica, aguas residuales, linajes, genómica, secuenciamiento.

ABSTRACT

Introduction: SARS-CoV-2 infection in humans results in viral RNA shedding primarily through respiratory droplets, but it can also be actively excreted via fecal matter. Wastewater pathogen surveillance has become a critical tool for epidemic preparedness, enabling the monitoring of emerging SARS-CoV-2 variants in populations with low clinical testing rates. For this reason, a passive sampling approach was implemented, consisting of using devices placed directly in the water flow for an extended period to collect a representative sample of the virus's behavior over time, without the need for active intervention or frequent collection. This approach was key in determining SARS-CoV-2 variants in Quito, Ecuador, during 2023-2024, with the goal of correlating these findings with national genomic surveillance data from hospital patients in Quito. **Methods:** A torpedo-type passive sampling device containing a nylon membrane, gauze, and a swab., the same one that was placed in wastewater collectors of the Quitumbe WWTP, for 24 hours a week. This method is a simpler approach to wastewater sampling than the conventional one that requires several liters of collection. Viral RNA was concentrated using PEG, followed by RNA extraction. SARS-CoV-2 detection was performed using real-time PCR, and multiplex PCR was used for sequencing with Oxford Nanopore Technology. Bioinformatic analysis was conducted using Freyja. **Results:** A predominance of the Omicron variant was observed, with lineages such as HN.1, EG.5.1.6, XBB.1.5.1.5, and HV.1 detected. Notably, the JN.1 lineage appeared during the infection peaks in Quito in October 2023 and February 2024, showing proportions similar to those found in circulating clinical samples from these periods and, in some cases, appearing even before clinical identification. It is important to highlight that the methodology used is highly sensitive to primer changes. Between March and July 2024, the results were inconsistent with clinical samples;. However, SARS-CoV-2 remained

detectable, only conserved sequences were identified, leading to incorrect lineage assignments, particularly of primitive lineages like JD.1. Starting in August 2024, recombinant lineages were detected, with the emergence of the KP.2.3 lineage in September, correlating with findings from human samples. **Conclusion:** Epidemiological surveillance of SARS-CoV-2 in wastewater provides a straightforward and effective method for genomic characterization of circulating variants and quantifying viral abundance in the population.

Keywords: Sars-CoV-2, passive torpedo-type sampling equipment, epidemiological surveillance, wastewater, lineages, genomics, sequencing.

CONTENT INDEX

INDEX OF TABLES

INDEX OF FIGURES

INTRODUCTION

SARS-CoV-2 belongs to the coronavirus family, characterized by its lipid envelope and a genomic structure consisting of positive-sense single-stranded RNA (+ssRNA). The SARS-CoV-2 genome is approximately 29.9 kilobases (kb) long and encodes several essential structural proteins, including the Spike (S) protein, the nucleocapsid (N) protein, the envelope (E) protein, and the membrane (M) protein. These structural components are critical for viral replication, infection, and interaction with the host immune system (Zhu et al., 2020).

Although small, the E protein is essential for virus morphogenesis and assembly. It plays a pivotal role in the formation of the viral particle and the release of new virions from infected cells. The M protein is responsible for the virion's shape and is crucial for virus assembly. It interacts with the S, E, and N proteins to facilitate the formation of new viral particles(Schoeman et al., 2020).

The N protein, an internal structural protein, directly binds to the viral RNA genome to form the nucleocapsid. This protein is vital for packaging the RNA within the virion and has immunomodulatory functions that interfere with the host's antiviral responses. Due to its high conservation and abundance, the N protein is a key target for diagnostic tests such as PCR-based assays (Surjit & Lal, 2008).

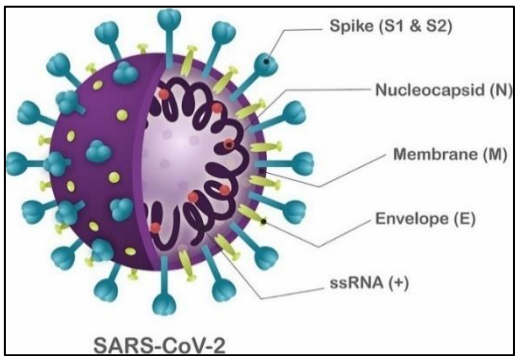

Figure 1 Structure Sars-CoV-2 (Santos et al., 2020) Antivirals Against Coronaviruses: Candidate Drugs for SARS-CoV-2 Treatment?. Frontiers. https://doi.org/10.3389/fmicb.2020.01818 Licencia CC BY 4.0.

The Spike (S) protein is the most studied component of the SARS-CoV-2 structure due to its crucial role in viral entry into host cells; it forms a trimeric complex consisting of three identical subunits that facilitate the binding and fusion process. These three copies of the S protein come together to create a stable trimer on the virus's surface. The trimer is heavily glycosylated, which helps shield the virus from immune recognition and enhances its ability to bind to the angiotensin-converting enzyme 2 (ACE2) receptor, thereby promoting the fusion of the viral and host cell membranes (Ramos-Montañez et al., 2008).

The Spike protein consists of three main subunits: **S1**, **S2**, and the **glycans**. The S1 subunit contains the receptor-binding domain (RBD), which is the specific region that interacts with the ACE2 receptor. The S2 subunit mediates the fusion between the viral and cellular membranes. Additionally, the glycans, which consist of protein regions covered in glycans, help the virus evade the host's immune response (Walls et al., 2020).

Mutations in the S protein are crucial in the emergence of new SARS-CoV-2 variants, such as the Delta and Omicron variants. These mutations can alter the protein's affinity for the ACE2 receptor, impacting the virus's transmissibility and its ability to evade immune detection and response (Harvey et al., 2021).

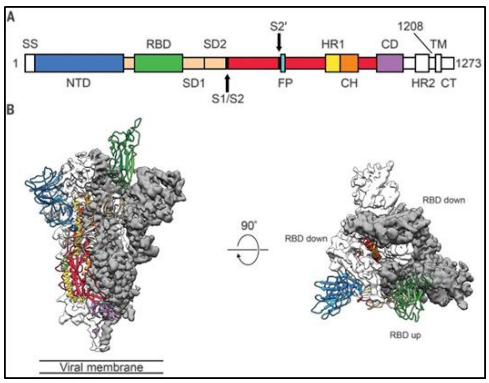

Figure 2 Structure of the spike protein (Wrapp et al., 2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. https://doi.org/10.1126/science.aax0902. Licencia CC BY 4.0.

Common Mutations in the Spike (S) Protein

Due to its critical role in viral entry, mutations in the S protein have been the focus of extensive research. The most notable mutations are described below:

D614G Mutation: This was one of the earliest mutations observed during the pandemic, first detected in March 2020. It involves the substitution of aspartic acid (D) with glycine (G) at position 614 of the Spike protein. This mutation increases the virus's infectivity by stabilizing the protein's open conformation, facilitating its interaction with the ACE2 receptor. As a result, the D614G mutation quickly became dominant and is now present in most SARS-CoV-2 lineages (Korber et al., 2020).

N501Y Mutation: This mutation changes asparagine (N) to tyrosine (Y) at position 501 in the Spike protein. It enhances the virus's binding affinity to the ACE2 receptor, potentially increasing its transmissibility. The N501Y mutation is present in several variants of concern, including Alpha (B.1.1.7), Beta (B.1.351), and Gamma (P.1) (Starr et al., 2020).

E484K Mutation: This mutation replaces glutamic acid (E) with lysine (K) at position 484 in the Spike protein. It has been linked to the virus's ability to evade neutralization by antibodies produced from prior infections or vaccination. A significant mutation is found in the Beta (B.1.351) and Gamma (P.1) variants (Greaney et al., 2021).

P681R Mutation: This mutation involves the substitution of proline (P) with arginine (R) at position 681, located near the furin cleavage site, a critical region for viral entry. It has been observed in the Delta variant (B.1.617.2). It is associated with increased efficiency in Spike protein cleavage, which could enhance the virus's ability to infect cells and improve its transmissibility (Cotten et al., n.d.).

K417N/T Mutation: This mutation changes lysine (K) to either asparagine (N) or threonine (T) at position 417 in the Spike protein. It is found in the Beta (B.1.351) and Gamma (P.1) variants and is associated with reduced antibody neutralization, suggesting possible immune evasion. However, its impact on ACE2 binding is moderate compared to other mutations in the receptor-binding domain (RBD) (Planas et al., 2021).

L452R Mutation: This mutation substitutes leucine (L) with arginine (R) at position 452 in the RBD of the Spike protein, it has been associated with increased binding affinity to the ACE2 receptor and partial resistance to some monoclonal antibodies, it is characteristic of the Delta (B.1.617.2) and Epsilon (B.1.427/B.1.429) variants (Motozono et al., 2021)

Country	Mutation rank									
	1	2	3	4	5	6	7	8	9	10
All*	D614G	P681H	N501Y	T716I	D1118H	A570D	S982A	A222V	L18F	S477N
USA	D614G	P681H	N501Y	T716I	A570D	D1118H	S982A	L452R	W152C	S13I
UK	D614G	P681H	N501Y	T716I	A570D	S982A	D1118H	A222V	L18F	L5F
France	D614G	N501Y	T716I	P681H	A570D	S982A	D1118H	S477N	A222V	E484K
Germany	D614G	N501Y	P681H	S982A	T716I	D1118H	A570D	A222V	L18F	S98F
Spain	D614G	P681H	T716I	N501Y	A570D	S982A	D1118H	A222V	D138Y	L18F
Italy	D614G	N501Y	P681H	T716I	D1118H	S982A	A570D	A222V	P272L	A262S
China	D614G	S12F	H49V	M153T	S50L	A688N	D1084E	Q498H	V1228I	F32S
India	D614G	P681R	E484Q	L452R	N440K	G142D	Q1071H	E154K	N501Y	Q677H
South Korea	D614G	N501Y	P681H	T716I	D1118H	A570D	S982A	L452R	S13I	W152C
Japan	D614G	M153T	Q675H	E484K	W152L	G769V	P681H	L54F	Q677H	G184S
Brazil	D614G	V1176F	E484K	N501Y	L18F	H655Y	P26S	D138Y	T20N	T1027I
South Africa	D614G	A701V	E484K	D80A	K417N	N501Y	D215G	L18F	R246I	A688V

Figure 3 Most common mutations in protein S (Negi et al., 2022). Regional and temporal coordinated mutation patterns in SARS-CoV-2 spike protein revealed by a clustering and network analysis. Nature. https://doi.org/10.1038/s41586-020-2012-7 Licencia CC BY 4.0

What is a Clade?

A clade refers to a group of virus variants that descend from a common ancestor, sharing specific key mutations. These groups are organized hierarchically based on the virus's evolutionary history. The Nextstrain system uses clades to monitor the global

evolution of SARS-CoV-2, providing a broader context for how the virus spreads and changes over time (Hadfield et al., 2018)

Other classifiers, such as GISAID, organize sequences into clades to monitor viral evolution (similar to Nextstrain). However, GISAID places a stronger emphasis on data sharing and global collaboration for real-time tracking of viral mutations and variants (GISAID, 2022).

The WHO Variant Classification system tracks the evolution of SARS-CoV-2 using a different approach. Instead of organizing variants into clades or lineages, the WHO groups them into **Variants of Concern (VOCs)** and **Variants of Interest (VOIs)** based on their potential impact on public health. This system prioritizes the level of risk posed by specific variants rather than their evolutionary relationships (World Health Organization, 2021).

What is a Lineage?

A lineage is a more specific subdivision within a clade, focusing on variants that share a particular set of more recent mutations. The PANGO system classifies SARS-CoV-2 lineages to track their transmission patterns. A lineage can represent a single branch within a clade and is often used for short-term epidemiological monitoring. For example, lineage B.1.1.7, which belongs to clade 20I (Alpha), contains a set of specific mutations that differentiate it from other lineages within the same clade (Rambaut et al., 2020).

Clades and lineages of SARS-CoV-2

Since its discovery in late 2019, the SARS-CoV-2 virus has undergone numerous mutations, leading to the emergence of different lineages and clades. These genetic variations are crucial for understanding the virus's evolution and play a key role in

epidemiological surveillance and assessing its clinical impact. Several nomenclature systems have been developed to classify the virus, with the PANGO and Nextstrain systems being among the most widely used.

The **PANGO system**, developed by (Rambaut et al., 2020), allows for a detailed classification of the virus into specific lineages that have evolved from ancestral ones. For instance, lineage B has given rise to several sublineages, such as B.1.1.7 (Alpha), B.1.351 (Beta), and B.1.617.2 (Delta). On the other hand, the Nextstrain system groups virus variants into clades, which represent sets of related lineages, providing a global perspective on the virus's evolution. Notable clades include 20A, 21K, and 22B, corresponding to the Alpha, Delta, and Omicron variants (Hadfield et al., 2018).

SARS-CoV-2 has evolved into numerous clades and lineages since the onset of the pandemic, with some having a more significant impact due to their increased transmissibility, immune evasion, or virulence. Below are some of the most substantial clades and lineages:

Clade 19A/19B: These were the first clades to emerge in Wuhan, China, in December 2019, and they form the foundation upon which subsequent variants have evolved. Clade 19A was responsible for the initial spread of the virus in Asia, while clade 19B quickly spread to other parts of the world (Worobey et al., 2020)

Clade 20A/20B/20C: These clades represent the second wave of variants that emerged globally in 2020. Lineage B.1 stood out as one of the most prevalent within these clades, particularly in Europe and North America. These lineages exhibited mutations in the Spike protein that facilitated viral entry into human cells through the ACE2 receptor (Candido et al., 2020).

Clade 21A (Delta): The Delta clade, corresponding to lineage B.1.617.2, was first identified in India in October 2020. It is notable for mutations in the Spike protein gene, such as L452R, which increases its affinity for the ACE2 receptor, and T478K, which enhances its ability to evade the immune response. Delta was responsible for large waves of infection in 2021 and demonstrated higher transmissibility and clinical severity than previous variants (Campbell et al., 2021).

Clade 21K (Omicron): The Omicron clade emerged in November 2021 and has been extensively studied due to its ability to evade immunity acquired from previous infections and vaccination. It is characterized by more than 30 mutations in the Spike protein, many of them in the receptor-binding domain (RBD), which help it evade neutralizing antibodies. Notable sublineages include BA.1, BA.2, BA.4, and BA.5, all exhibiting similar epidemiological behavior (Tegally et al., 2022).

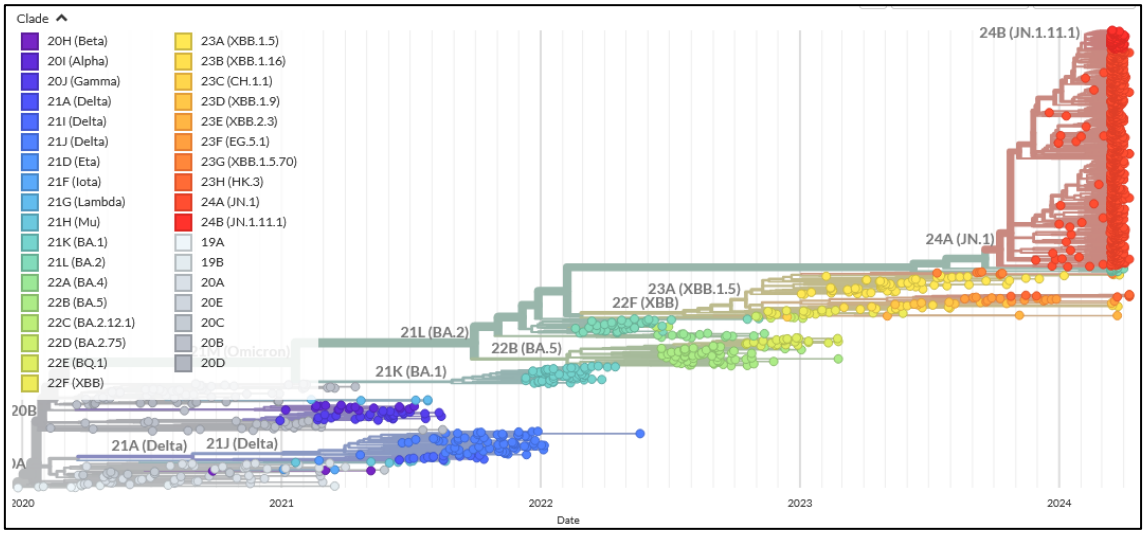

Figure 4 Clades and lineages Sars-CoV-2 (Nextstrain, 2022).

Virus Detection

The detection of SARS-CoV-2 has been a crucial component in combating the pandemic, enabling the rapid and accurate identification of infected individuals, monitoring the spread of the virus, and assessing its evolution over time. The primary

detection techniques focus on analyzing the virus's genetic material and the host's immune response; among the most used methods are:

RT-qPCR: This technique detects viral RNA in respiratory samples, such as nasopharyngeal swabs, by amplifying specific regions of the viral genome, usually targeting the N, E, or S genes. It is susceptible and specific, even allowing for virus detection in the early stages of infection when viral load is low. The process involves converting viral RNA into complementary DNA (cDNA) using reverse transcriptase, followed by amplification through PCR. If viral cDNA is present in the sample, a fluorescent signal is emitted, confirming the presence of the virus (Corman et al., 2020).

Antigen Tests: These tests rapidly detect SARS-CoV-2 by identifying viral proteins in respiratory samples, such as the nucleocapsid. They are faster and less expensive than RT-PCR, providing results in approximately 15-30 minutes. However, antigen tests are less sensitive, especially in individuals with a low viral load or during the early stages of infection (Scohy et al., 2020).

Serological Tests: These tests detect antibodies against SARS-CoV-2 in the blood, indicating whether a person has been previously exposed to the virus, even if they are asymptomatic or have recovered. The most common tests measure IgM and IgG antibodies. While useful for seroprevalence studies and assessing population immunity, they are not suitable for diagnosing active infections (Long et al., 2020),

Whole Genome Sequencing: This method provides a detailed analysis of the virus's genetic structure, aiding in detecting and monitoring viral evolution and identifying emerging variants. The most widely used technologies for this purpose include next-generation sequencing (NGS) platforms like Illumina and Oxford Nanopore. The process begins with converting viral RNA into cDNA, then fragmented and

sequenced into millions of small reads. These fragments are assembled to reconstruct the complete viral genome, allowing for the identification of specific mutations and the classification of viral lineages. This technique is critical for epidemiological surveillance and understanding the spread of variants, as it enables population-level monitoring and response to new variants that may impact transmissibility or vaccine efficacy (Oude Munnink et al., 2020).

The correct use of bioinformatics tools is essential for accurately assigning viral lineages, a key element in epidemiological studies of SARS-CoV-2. These tools efficiently analyze and classify viral sequences, detect emerging variants, and comprehensively view their evolution in the population. Some of the most important tools include:

EPI2ME™: Developed by Oxford Nanopore Technologies, this bioinformatics platform analyzes high-precision sequencing data in real time. It facilitates the annotation of viral genomes and sequencing data analysis through automated and customizable workflows (Oxford Nanopore Technologies, 2020).

Nextclade: Created by the Nextstrain team, this tool classifies viral sequences and compares them with a global reference. It provides information on the mutations in a SARS-CoV-2 sequence and classifies them into lineages based on their phylogenetic relationships. It also identifies relevant clades and variations while assessing the quality of the analyzed sequences (Aksamentov et al., 2021).

FREYJA: This specialized bioinformatics tool identifies viral lineages by analyzing viral sequences obtained from wastewater samples. It is based on detecting viral variants in the population and accurately estimates the prevalence of SARS-CoV-2 lineages in different communities (Karthikeyan et al., 2022).

Importance of Epidemiological Surveillance in Clinical Samples

Epidemiological surveillance in clinical samples is essential for disease control and prevention, particularly during pandemics like the one caused by SARS-CoV-2. This type of surveillance enables the detection of circulating pathogens in the population by monitoring biological samples such as blood, nasopharyngeal swabs, and other body fluids (CDC, 2020). The significance of this process lies in its ability to identify specific viral variants or lineages, allowing researchers to evaluate the dynamics of viral transmission and mutation. Early detection of viral variants through this type of monitoring facilitates the correct assignment of lineages, which is crucial for the timely implementation of public health measures. This is particularly relevant when detecting key mutations in essential genes, such as the one encoding the Spike protein, as these mutations can directly affect the effectiveness of clinical interventions and the overall dynamics of the pandemic (Harvey et al., 2021).

Virus Detection in Wastewater and Its Importance

The detection of SARS-CoV-2 in wastewater has become a critical tool in epidemiological surveillance, especially for monitoring the spread of the virus across large populations. This method, known as wastewater surveillance, enables the identification of viral RNA fragments excreted in human waste, providing an indirect yet highly effective means to estimate the circulation of the virus within a community (Schang et al., 2021).

One of the major advantages of wastewater surveillance is its capacity to monitor large population groups without requiring individual testing, which makes it both efficient and non-invasive, it is particularly valuable in areas where access to clinical testing may be limited. Additionally, it has been successfully implemented in various settings,

including university campuses, communities, and even entire regions, to anticipate potential outbreaks before a significant increase in clinical cases is observed (Ahmed et al., 2020).

Asymptomatic individuals who do not show symptoms and do not seek medical attention or undergo testing play a crucial role in this type of surveillance, wastewater surveillance captures the viral spread among these asymptomatic individuals, who would otherwise remain undetected by traditional clinical testing methods, thus, it provides a more accurate view of the true extent of viral circulation within a community (Kitamura et al., 2021).

It is essential to highlight that the virus can be detected in wastewater even before individuals exhibit symptoms, this early detection acts as an advanced warning system, alerting authorities to the presence of SARS-CoV-2 in a community, thereby enabling the prompt implementation of control and mitigation measures (Wurtzer et al., 2020). Furthermore, sequencing viral RNA from wastewater allows for the virus's detection and the identification of circulating SARS-CoV-2 variants in each region. This is essential for tracking the emergence of new variants that may be more transmissible or capable of evading the immunity provided by vaccines (Crits-Christoph et al., 2021). Lastly, wastewater surveillance can reveal trends in infection rates within a community, helping to track increases or decreases in viral spread. These data are beneficial for evaluating the effectiveness of public health policies or vaccination campaigns (Peccia et al., 2020).

Current Status of Epidemiological Surveillance in Wastewater

Globally, wastewater-based surveillance has greatly enhanced the early detection and tracking of emerging variants of various microorganisms, further supported by genomic sequencing. International efforts, such as the Global Influenza Surveillance and

Response System (GISRS) and its extension to include SARS-CoV-2, have facilitated the exchange of information on viral lineages and variants among countries. Additionally, GISAID, an international database, has played a crucial role in sharing viral sequences, aiding in understanding how and where variants of concern (VOC) are circulating (World Health Organization, 2021).

Numerous countries worldwide have adopted wastewater surveillance as a key epidemiological control tool to detect SARS-CoV-2, this method has been implemented across various regions, including Europe, North America, Asia, and Oceania. In the United States, the Centers for Disease Control and Prevention (CDC), through the National Wastewater Surveillance System (NWSS), have developed a monitoring framework that covers approximately 80% of the country's population. This system enables the detection of SARS-CoV-2 in over 400 sites nationwide, providing authorities with data to support rapid decision-making (CDC, 2020). Similar efforts have been made in the Netherlands, the United Kingdom, Spain, Australia, Japan, Canada, and New Zealand.

In Latin America, several countries, including Brazil, Argentina, Mexico, Colombia, and Chile, have also implemented epidemiological surveillance of SARS-CoV-2 in wastewater as a complementary tool to monitor viral spread. These countries have utilized this technique to gather data that supplement the information obtained from clinical testing, especially in identifying asymptomatic infections, which are not always captured by traditional clinical diagnostic systems (Centers for Disease Control and Prevention, 2024).

This proactive approach has enabled authorities to make informed decisions regarding public health measures. However, one significant challenge in the region has been the infrastructure of sanitation systems. In some areas, inadequate sewer coverage

limits the reach of wastewater surveillance, particularly in rural or marginalized communities (Zhang et al., 2022).

Cryptic Lineages

Ongoing global monitoring of SARS-CoV-2 has underscored the importance of wastewater surveillance as a tool for detecting and tracking viral variants. While clinical testing remains the primary method for identifying circulating lineages, there are cases where variants are found in wastewater but not detected in clinical samples. These variants, often referred to as cryptic lineages, are of particular interest because they represent viral populations circulating among underdiagnosed, unmonitored groups, or possibly even in non-human host (Smyth et al., 2022).

The detection of such lineages in wastewater emphasizes the value of environmental surveillance in capturing viral diversity that might otherwise go unnoticed through traditional clinical testing. Wastewater monitoring enables us to track viral evolution and the emergence of new variants, providing critical insights into reservoirs of viral persistence and transmission (Knight, 2022).

Cryptic SARS-CoV-2 lineages exhibit distinct genetic profiles that differ from known variants and may arise due to selective pressures, such as the immune response in specific environments (Callaway, 2022).

Studies conducted in New York have identified several cryptic lineages in wastewater that do not share a close common ancestor with the predominant variants in circulation, these lineages display unique mutations in regions of the viral genome that are not typically found in the known SARS-CoV-2 variants circulating in humans, some of these mutations, such as A23056C and C24044T in the Spike protein gene, are rare in

clinical samples, suggesting that these lineages could be evolving in specific environments, such as sewage systems (Gregory et al., 2022).

Cryptic lineages have shown signs of evolution under positive selection, indicating that they are adapting to new conditions, potentially in response to immune system pressures. For example, in Missouri, a cryptic lineage was observed with specific mutations in the Spike protein, known to contribute to immune evasion, while these mutations are similar to those found in the Omicron variant, these cryptic lineages were detected prior to Omicron's emergence, suggesting an independent yet convergent evolution (Gregory et al., 2023).

The detection of cryptic lineages presents significant challenges, as many of these lineages have not been identified in clinical samples, and their tracking requires specialized sequencing techniques, nonetheless, cryptic lineages can offer valuable insights into viral evolution, particularly in non-human reservoirs or under-monitored populations, underscoring the importance of genomic surveillance in wastewater as a key epidemiological tool (Gregory et al., 2022).

GENERAL OBJECTIVE

To analyze the SARS-CoV-2 viral lineages present in wastewater using bioinformatics tools, to monitor variant circulation and provide relevant data for epidemiological surveillance in Ecuador.

SPECIFIC OBJECTIVES

- Implement sequencing techniques and bioinformatic analyses to detect viral lineages in wastewater samples from the Quitumbe treatment plant from October 2023 to September 2024, from 7am on Monday to 7am on Tuesday, that is, for a period of 24 hours.

- To characterize the circulating SARS-CoV-2 variants in the samples by assigning lineages using tools such as FREYJA, Nextclade, and EPI2ME™.
- To identify and analyze the presence of emerging lineages to evaluate their transmission dynamics and possible relevant mutations.
- To contribute to epidemiological surveillance in Ecuador by integrating the results into international open databases like GISAID, facilitating lineage comparison with other regions.

MATERIALS AND METHODS

Wastewater samples were collected from October 2023 to September 2024, including two seasons (rainy and dry). We followed the protocol by Schang and colleagues (2021) [OBJ] and used a passive 3D-printed device for sample collection (see **Figure 5**).

Due to its torpedo-like design, the passive sampler prevents large debris from adhering to the equipment and obstructing the entry of wastewater. The device was equipped with a nylon membrane (0.45 µm pore size), gauze, cotton swabs, and a metal rod to ensure that it remained fully submerged in the wastewater, thus ensuring the collection of a representative sample. This device is relatively easy to deploy in various locations and more convenient to transport than traditional wastewater sampling methods.

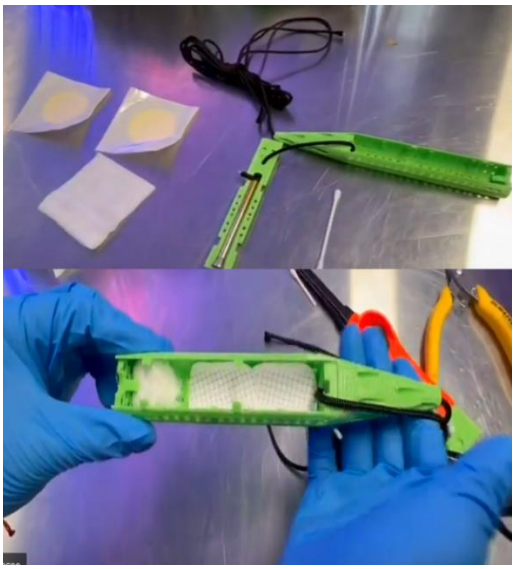

Figure 5 Passive sampling 3D printed device (Schang et al., 2021)

Sampling standardization

The device was firstly tested for 3 hours in the San Pedro River (see **Figure 6**), specifically along the El Chaquiñán Ecological Route, during the peak sewer service usage hours, that is, from 11am to 2pm. However, during sample processing, it was observed that RNA was not being adequately concentrated. As a result, the deployment time was extended to 24 hours, which provided better results.

Figure 6 San Pedro River, location of the 3D printed device to monitor SARS-CoV-2 in waste discharge.

Sample collection

After standardizing the sampling process, the devices were installed at the Quitumbe treatment plant to ensure greater control and safety during placement, this plant

serves 11 neighborhoods in the Metropolitan District of Quito, which are: Manuelita Sáenz, San Alfonso, Nuevos Horizontes, Los Cóndores, Los Arrayanes, San Francisco del Sur, Martha Bucaram de Roldós, La Ecuatoriana, Las Orquídeas, La Concordia y Ninallacta (see **Figure 7**), making the investigation more targeted by detecting circulating lineages specific to that population. The samplers were immersed for 24 hours once a week in the incoming wastewater at the Quitumbe WWTP, where the flow rate is 75 L/s. Before any treatment was applied, a single weekly sample was taken, resulting in a total of 47 samples throughout the study. To prevent large debris from adhering to and obstructing the device, the samplers were placed in black mesh bags, after the sampling period, the devices were retrieved and transported in a cooler at 4°C to the Microbiology Institute at Universidad San Francisco de Quito. In a laminar flow hood, the devices were carefully disassembled, and the membrane, swab, and gauze were placed in separate sterile Ziploc bags for further processing.

Figure 7 Neighborhoods connected to the Quitumbe Wastewater Treatment Plant.

RNA concentration

Because SARS-CoV-2 is present in low concentrations in wastewater samples, the RNA must first be concentrated to ensure sufficient viral material for detection and

analysis, by preparing a 2X PEG solution, following the protocol by Woods and Jones (2021), in this protocol, 20g of PEG and 3.48g of NaCl are added to a BOECO flask, filled with 200mL of nuclease-free water, and autoclaved. Following the method by Schang et al. (2021), about 4g of the sample is placed in a sterile bag between the gauze, membrane, and swab, then, 10mL of 10X dPBS and 0.05% Tween 20 are added, and the samples are mixed in a Stomacher for 2 minutes at 200rpm. The supernatant is transferred to a 10mL tube and centrifuged at 400rpm for 10 minutes. After centrifugation, approximately 9mL of the supernatant is transferred, and 20% of the 2X PEG solution is added, the tubes are incubated for 24 hours at 4°C, and finally, the samples are centrifuged again at 400rpm for 15 minutes, the supernatant is discarded, and the remaining pellet is resuspended in 300µL of Shield2X.

It should be noted that during the RNA concentration process the samples (gauze, membrane, and swab) were initially tested separately. However, due to the low concentrations obtained from the individual samples, it was decided to pool the samples collected on the same date, for later the final RNA concentration analysis.

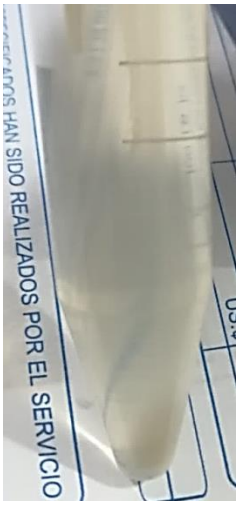

Figure 8 Pellet prepared for elution in Shield 2X.

RNA extraction

During the extraction process, several kits were tested to determine which yielded the highest RNA concentrations. Initially, the QIAamp Fast DNA Stool Mini Kit was

used, another extraction method was the Zymo DNA/RNA Miniprep Kit and finally the samples were extracted using the Zymobiomics Quick-RNA Viral Kit, with some modifications to the standard procedure. First, 400µL of Viral Buffer RNA was added to 200µL of the sample and vortexed. The mixture was then transferred to the ZYMO Spin IC Column and centrifuged for 2 minutes at 13,000 rpm. The column was then transferred to a clean tube. Next, 500µL of Viral Wash Buffer was added to the column and centrifuged for 2 minutes at 13,000 rpm. The liquid from the tube was discarded, and the column was preserved. This step was repeated. Then, 500µL of ethanol (95-100%) was added to the column and centrifuged for 1 minute at 13,000 rpm. The columns were carefully removed and placed in sterile 1.5mL Eppendorf tubes. Finally, 20µL of DNase and RNase-free water was added to the column and centrifuged for 1 minute. It is recommended to add 10µL first, then the other 10µL, and incubate at room temperature for 3 to 5 minutes before centrifugation. It should be noted that positive and negative controls were used for all samples to verify the effectiveness of the kit and to rule out any potential contamination and the RNA was then quantified using Nanodrop or Qubit.

Molecular identification of SARS-CoV-2

Diagnosis of SARS-CoV-2 or Real-time PCR in the identification of the positive samples for Sars CoV-2, the Luna SARS-CoV-2 RT-qPCR Multiplex Assay Protocol (NEW ENGLAND Biolabs, Ipswich, MA) was used following the protocol, the PCR mixture consisted of a final volume of 20µL and contained 11µL of Nuclease-free water, 5µL of Luna Probe One-Step RT-qPCR 4X Mix with UDG, 2µL of SARS-CoV-2 Primer/Probe Mix (N1/N2/RP) (10X) and 2µL of samples RNA. The thermocycling procedure was conducted in a thermocycler (Bio-Rad Laboratories, Inc, California, USA) with carryover prevention at 25°C for 30 seconds, then reverse transcription at 55°C for 10 minutes, an initial denaturation at 95°C for 1 minute, followed by 45 cycles of

denaturation at 95°C for 10 seconds and a final extension of 60°C for 30 seconds. A multiplex PCR was conducted following the NEBNext® ARTIC SARS-CoV-2 RT-PCR Module protocol outlined by (NEW ENGLAND Biolabs, Ipswich, MA) with slight modifications. The cDNA is prepared by placing 2 µL of LunaScript® RT SuperMix (5X) and 8µL of RNA. The thermocycling procedure was conducted in a thermocycler (Thermo Fisher Scientific, Massachusetts, USA) with primer annealing at 25ºC for 2 minutes, a cDNA synthesis at 55ºC for 20 minutes, and a heat inactivation at 95ºC for 1 minute. The PCR mix had a final volume of 12.5µL and contained 6.25µL of Q5® Hot Start High-Fidelity 2X Master Mix, 1.25µL each of NEBNext ARTIC SARS-CoV-2 Primer Mix 1 and NEBNext ARTIC SARS-CoV-2 Primer Mix 2. PCR reactions from Pool A and Pool B were combined for each sample, and 5µL of cDNA was added, the thermocycling procedure was conducted in a thermocycler (Thermo Fisher Scientific, Massachusetts, USA) with an initial denaturation at 98ºC for 30 seconds followed by 35 cycles of denaturation at 95°C for 15 seconds and a final extension of 63°C for 5 minutes. It should be noted that positive and negative controls were used for all samples to verify the effectiveness of the kit and to rule out any potential contamination. As a result of running the 1.5% agarose gel, faint bands and primer dimers were observed.

Since the NEBNext® Varskip Short ARTIC SARS-CoV-2 primers were not available, nCoV-2019 (LoCost) V.3 primers from New England Biolabs were used instead. The cDNA preparation remained unchanged; However, the preparation of the mix for the multiplex PCR was adjusted. A final volume of 25µL, consisting of 12.5µL of Q5® Hot Start High-Fidelity 2X Master Mix, 5µL of nuclease-free water, and 3.5µL of nCoV-2019 (LoCost) V.3 Primer Mix 1 and Primer Mix 2. PCR reactions from groups A and B were combined for each sample and 4µL of cDNA was added. The thermocycling procedure was conducted in a thermocycler (Thermo Fisher Scientific, Massachusetts, USA). The

hot start was at 98ºC for 30 seconds, followed by 35 cycles of denaturation at 98°C for 15 seconds and a final extension of 65°C for 5 minutes.

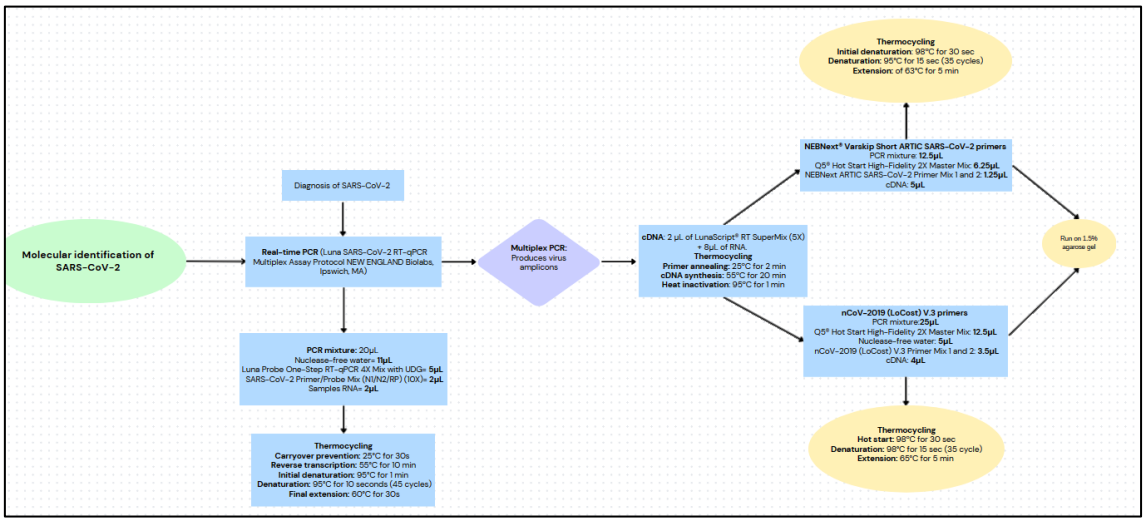

Figure 9 Flow chart of molecular identification Sars-CoV-2

Amplicon sequencing analysis

Following the amplification of PCR products for SARS-CoV-2, sequencing was performed on amplicons that had been diluted at a 1:4 ratio, for these samples, 15µL of nuclease-free water was added, and 5µL was distributed between Pool A and Pool B. Sequencing was also conducted on undiluted amplicons, which were sequenced using R10.4.1 flow cells from Oxford Nanopore technology, following the methodology described by Oxford Nanopore Technologies (2021) with slight modifications.

In accordance with the ligation sequencing amplicons native barcoding protocol, the previously diluted amplicons were combined with 2.5µL of end-prep, 5µL of Blunt/TA Ligase Master Mix, and 1.25µL of Native Barcode, the mixture was incubated using a thermal cycler at 20°C for 5 minutes and at 65°C for another 5 minutes, the reaction was then stopped by adding 1µL of EDTA, all the samples were mixed in a 1.5mL Eppendorf tube, and 0.4X Ampure XP beads were added, the samples were incubated for 10 minutes at room temperature on the magnetic rack, the pellet was washed with 80% ethanol, and then it was resuspended in 30µL of nuclease-free water, followed

by another 10-minute incubation at room temperature. After that, 3µL of Native Adapter, 10µL of NEBNext Quick Ligation Reaction Buffer (5X), and 5µL of Quick T4 DNA Ligase were added, the mixture was incubated for 20 minutes at room temperature, then, 20µL of AMPure XP Beads were added and the sample was incubated on a Hula mixer for 10 minutes at room temperature. It was placed on the magnetic rack for 5 minutes, the supernatant was removed, and the pellet was resuspended in 125µL of Short Fragment Buffer (SFB). The process was repeated on the magnetic rack, after which the supernatant was removed and the pellet was resuspended in 15µL of Elution Buffer (EB), the sample was incubated for 10 minutes at room temperature, and then 1µL containing the DNA library was transferred to a clean tube and quantified using Qubit.

For priming and loading the SpotON flow cell, 1100µL of Flow Cell Flush (FCF), 25µL of Flow Cell Tether (FCT), and 5µL of Bovine Serum Albumin (BSA) at 50 mg/mL were mixed in an Eppendorf tube. A total of 800µL was placed in the priming pore, ensuring no air bubbles were present, in another tube, the library was prepared for loading into the flow cell by adding 37.5µL of Sequencing Buffer (SB), 25.5µL of Library Beads (LIB), and 12µL of the DNA library, a total of 200µL of the priming mix was loaded into the Flow Cell priming port (not the SpotON sample port), avoiding air bubbles, and 75µL of the prepared library was added to the Flow Cell via the SpotON sample port dropwise. Each drop was allowed to flow into the port before adding the next. Finally, the Flow Cell was placed in the GridION (Oxford Nanopore Technology, Oxford, UK) and sequenced for approximately 24 hours.

Bioinformatics analysis

The files generated after sequencing are analyzed using the bioinformatics tools EPI2ME™ and FREYJA. These tools create a consensus sequence, assign lineages to each sample, and produce graphs showing the relative abundance of each lineage present.

This enables observation of changes or circulation of lineages at different times of the year in a specific location. Additionally, the sequencing data were analyzed using Nextclade and were compared with those identified by FREYJA to assess data concordance.

RESULTS

Part 1: Detection of SARS-CoV-2 in wastewater protocol standardization

After conducting multiple tests and comparing the results from Nanodrop, Qubit, and qPCR, the highest RNA concentrations were obtained by pre-concentrating the sample using the 2X PEG preparation method outlined by Woods & Jones (2021), followed by the RNA concentration protocol described by Schang et al. (2021). For RNA extraction, the Zymobiomics Quick-RNA Viral Kit was chosen due to its availability, reliable results, and the benefit of a relatively short processing time.

Table 1. RNA Concentration and Quality of Quitumbe WWTP Samples Measured by NanoDrop.

Concentration and quality in Nanodrop							
Date	**Sample**	**Quick RNA viral kit zymo research**			**Quick RNA viral kit zymo research+PEG**		
		ng/uL	**A260/A280**	**A260/A230**	**ng/uL**	**A260/A280**	**A260/A230**
30 October 2023	**WW1**	46,2	2,16	1,26	30,8	1,64	0,57
06 November 2023	**WW2**	3,3	1,6	0,06	31,8	1,54	0,55
13 November 2023	**WW3**	-8,1	1,74	0,92	66,4	1,67	0,82
20 November 2023	**WW4**	-1,8	2,37	0,48	61,4	1,52	0,61
27 November 2023	**WW5**	3	-0,8	-0,32	34,9	1,62	0,76
27 November 2023 (Cárcamo)	**WW6**	198,4	2,13	1,64	64,1	1,65	0,84
04 December 2023 (Cárcamo)	**WW7**	61	2,11	1,58	90,4	1,69	0,88
05 December 2023	**WW8**	81,8	2,09	2,07	97,1	1,79	1,33
11 December 2023	**WW9**	-6,2	1,58	0,18	67,7	1,81	1,58
18 December 2023	**WW10**	6,4	1,98	0,03	106,1	1,74	1,14

The RNA quantification results using the Quick RNA viral kit zymo research alone and the Quick RNA viral kit zymo research with PEG showed notable differences. The PEG method consistently yielded

higher RNA concentrations, suggesting that pre-concentration improves the detection of viral RNA in wastewater samples. Moreover, the A260/280 and A260/230 ratios, which are indicators of RNA purity, were generally within acceptable ranges for both methods. However, slight variations in purity ratios suggest potential differences in sample composition or contaminants between the two approaches.

Table 2. Comparison of Ct Values for RNA Extracted Using the Quick RNA viral kit zymo research and Quick RNA viral kit zymo research with PEG

Ct value of the samples					
Date	**Sample**	**Quick RNA viral kit zymo research**		**Quick RNA viral kit zymo research+PEG**	
		Fluorophore FAM (Gene N1)	**Fluorophore HEX (Gene N2)**	**Fluorophore FAM (Gene N1)**	**Fluorophore HEX (Gene N2)**
30 October 2023	**WW1**	35,62	35,86	36,87	37,41
06 November 2023	**WW2**	45	42,9	37,23	37,23
13 November 2023	**WW3**	45	38,01	36,47	38,3
20 November 2023	**WW4**	40,8	38,97	35,6	36,19
27 November 2023	**WW5**	36,1	35,48	34,91	35,5
27 November 2023 (Cárcamo)	**WW6**	44,81	36,62	35,43	45
04 December 2023 (Cárcamo)	**WW7**	45	38,27	33,23	35,47
05 December 2023	**WW8**	45	40	34,16	34,89
11 December 2023	**WW9**	45	41,09	33,51	35,3
18 December 2023	**WW10**	36,46	N/A	35,51	36,42

The comparison of Ct values between the Quick RNA viral kit zymo research alone and the Quick RNA viral kit zymo research with PEG shows that the PEG-treated samples consistently have slightly early Ct values. Since early Ct values indicate a higher concentration of detectable RNA, this suggests that the PEG pre-concentration method improves the recovery and detection of viral RNA from wastewater. Therefore, enrichment the samples with PEG enhances the sensitivity of the assay, making it a more effective approach for detecting SARS-CoV-2 in wastewater samples with lower RNA concentrations.

For qPCR, artic (New England Biolabs, Ipswich, MA) was the only kit tested, as it demonstrated effective detection and amplification of the virus's N1 and N2 genes, these genes are crucial for confirming the presence of SARS-CoV-2, and their CT values provide insights into viral circulation trends at the study site, supporting epidemiological surveillance and contributing to effective virus control efforts.

Regarding multiplex PCR, after multiple rounds of standardization and primer adjustments based on sequencing results, the oxford set proved to be the most effective, offering improved lineage assignment and broader viral genome coverage.

For sequencing, optimal lineage assignment was achieved by diluting the samples at a 1:4 ratio before beginning the procedure; The bioinformatics analysis revealed consistent detection of similar lineages with both the FREYJA and NEXTCLADE tools, confirming accurate lineage assignment for each sample.

Part 2: Epidemiology of SARS-CoV-2 in wastewater between 2023 -2024

qPCR quantifications

As shown in **Figure 10** and **Table 1**, the N1 and N2 genes were detected in all samples, confirming the circulation of SARS-CoV-2 in this area of Quito. However, the observed CT values were relatively late, likely due to viral RNA fragmentation in the samples. This may hinder an accurate determination of viral concentration, unlike what is typically seen in clinical samples. **Figure 10** also shows peaks in detecting the N1 gene from October 2023 to January 2024. During approximately two weeks, both the N1 and N2 genes exhibited late CT values. After that, on February 13, 2024, peaks for both genes reappeared. Notably, from March 2024, intermittent peaks were observed exclusively for the N2 gene, which persisted until the end of the study. It is worth noting that the N2 gene, marked with the HEX fluorophore, showed slightly earlier CT values, suggesting that factors such as RNA degradation or genetic mutations may make the N2 gene easier to quantify or indicate a higher viral load.

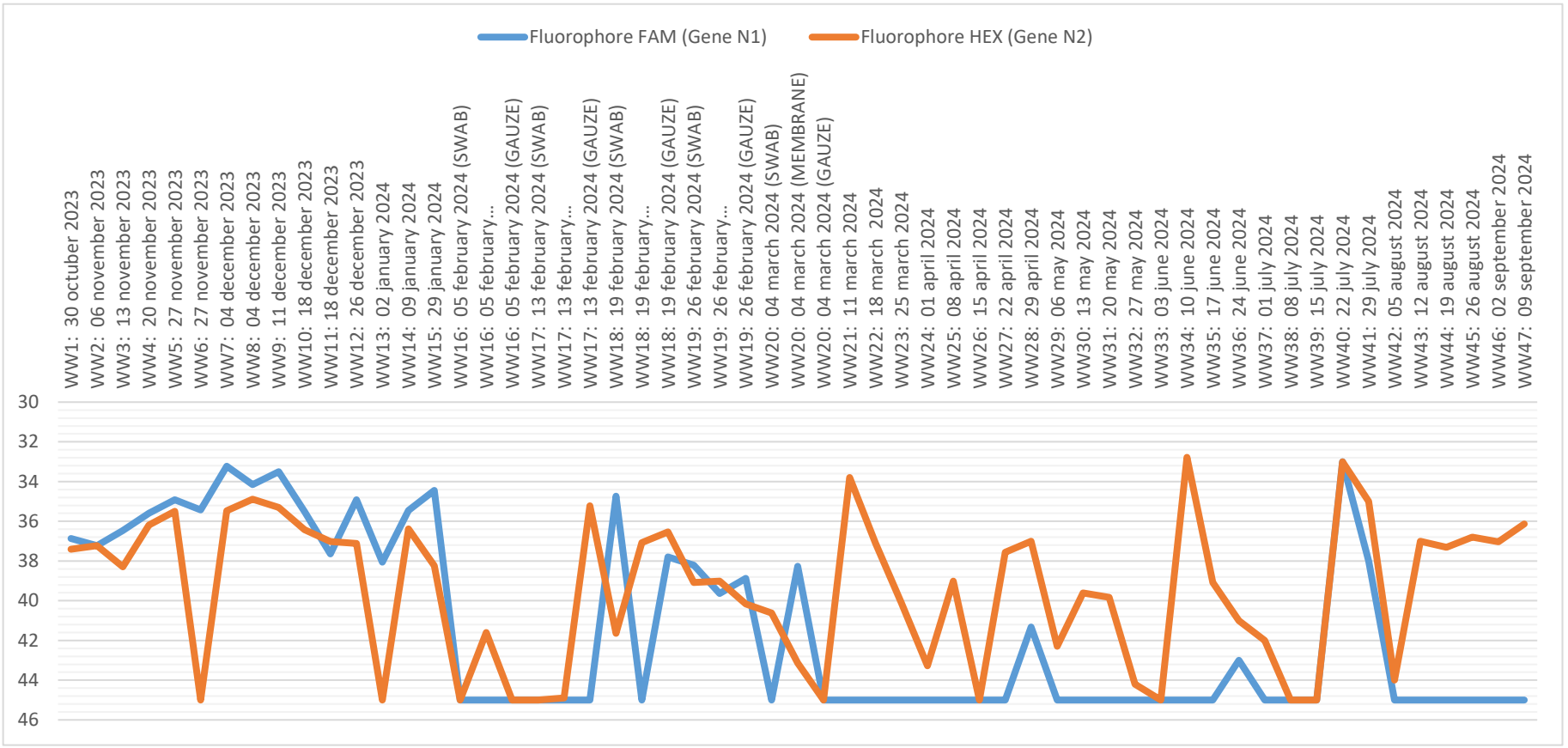

Figure 10 Sars CoV-2 cycle threshold of wastewater samples

The wastewater samples were collected over 11 months. The CT values for the N1 gene, marked with the blue FAM fluorophore, and the N2 gene, marked with the orange HEX fluorophore, were analyzed. The data show a consistent late trend in detecting the virus across the samples.

Table 3. Types of samples used for qPCR with their corresponding CT value.

Sample	Location	Swab	Membrane	Gauze	Date	Fluorophore FAM (Gene N1)	Fluorophore HEX (Gene N2)
WW1	Entrance WWTP			YES	**WW1: 30 octuber 2023**	36,87	37,41
WW2	Entrance WWTP			YES	**WW2: 06 november 2023**	37,23	37,23
WW3	Entrance WWTP			YES	**WW3: 13 november 2023**	36,47	38,3
WW4	Cárcamo WWTP			YES	**WW4: 20 november 2023**	35,6	36,19
WW5	Cárcamo WWTP			YES	**WW5: 27 november 2023**	34,91	35,5
WW6	Entrance WWTP			YES	**WW6: 27 november 2023**	35,43	45
WW7	Cárcamo PTAR			YES	**WW7: 04 december 2023**	33,23	35,47
WW8	Entrance WWTP			YES	**WW8: 04 december 2023**	34,16	34,89
WW9	Entrance WWTP			YES	**WW9: 11 december 2023**	33,51	35,3
WW10	Entrance WWTP			YES	**WW10: 18 december 2023**	35,51	36,42
WW11	San Pedro river			YES	**WW11: 18 december 2023**	37,64	37,02
WW12	Entrance WWTP			YES	**WW12: 26 december 2023**	34,91	37,11
WW13	Entrance WWTP			YES	**WW13: 02 january 2024**	38,06	45
WW14	Entrance WWTP			YES	**WW14: 09 january 2024**	35,46	36,39
WW15	Entrance WWTP			YES	**WW15: 29 january 2024**	34,45	38,24
WW16	Entrance WWTP	YES			**WW16: 05 february 2024 (SWAB)**	45	45
WW16	Entrance WWTP		YES		**WW16: 05 february 2024(MEMBRANE)**	45	41,6
WW16	Entrance WWTP			YES	**WW16: 05 february 2024 (GAUZE)**	45	45
WW17	Entrance WWTP	YES			**WW17: 13 february 2024 (SWAB)**	45	45
WW17	Entrance WWTP		YES		**WW17: 13 february 2024(MEMBRANE)**	45	44,9
WW17	Entrance WWTP			YES	**WW17: 13 february 2024 (GAUZE)**	45	35,22
WW18	Entrance WWTP	YES			**WW18: 19 february 2024 (SWAB)**	34,73	41,64

WW18	Entrance WWTP		YES		**WW18: 19 february 2024(MEMBRANE)**	45	37,07
WW18	Entrance WWTP			YES	**WW18: 19 february 2024 (GAUZE)**	37,79	36,53
WW19	Entrance WWTP	YES			**WW19: 26 february 2024 (SWAB)**	38,21	39,08
WW19	Entrance WWTP		YES		**WW19: 26 february 2024(MEMBRANE)**	39,64	39,01
WW19	Entrance WWTP			YES	**WW19: 26 february 2024 (GAUZE)**	38,88	40,16
WW20	Entrance WWTP	YES			**WW20: 04 march 2024 (SWAB)**	45	40,62
WW20	Entrance WWTP		YES		**WW20: 04 march 2024 (MEMBRANE)**	38,27	43,14
WW20	Entrance WWTP	YES	YES	YES	**WW20: 04 march 2024 (GAUZE)**	45	45
WW21	Entrance WWTP	YES	YES	YES	**WW21: 11 march 2024**	45	33,8
WW22	Entrance WWTP	YES	YES	YES	**WW22: 18 march 2024**	45	37,1
WW23	Entrance WWTP	YES	YES	YES	**WW23: 25 march 2024**	45	40,1
WW24	Entrance WWTP	YES	YES	YES	**WW24: 01 april 2024**	45	43,27
WW25	Entrance WWTP	YES	YES	YES	**WW25: 08 april 2024**	45	39,01
WW26	Entrance WWTP	YES	YES	YES	**WW26: 15 april 2024**	45	45
WW27	Entrance WWTP	YES	YES	YES	**WW27: 22 april 2024**	45	37,56
WW28	Entrance WWTP	YES	YES	YES	**WW28: 29 april 2024**	41,32	37,01
WW29	Entrance WWTP	YES	YES	YES	**WW29: 06 may 2024**	45	42,3
WW30	Entrance WWTP	YES	YES	YES	**WW30: 13 may 2024**	45	39,61
WW31	Entrance WWTP	YES	YES	YES	**WW31: 20 may 2024**	45	39,83
WW32	Entrance WWTP	YES	YES	YES	**WW32: 27 may 2024**	45	44,21
WW33	Entrance WWTP	YES	YES	YES	**WW33: 03 june 2024**	45	45
WW34	Entrance WWTP	YES	YES	YES	**WW34: 10 june 2024**	45	32,78
WW35	Entrance WWTP	YES	YES	YES	**WW35: 17 june 2024**	45	39,08
WW36	Entrance WWTP	YES	YES	YES	**WW36: 24 june 2024**	43	41
WW37	Entrance WWTP	YES	YES	YES	**WW37: 01 july 2024**	45	42

WW38	Entrance WWTP	YES	YES	YES	**WW38: 08 july 2024**	45	45
WW39	Entrance WWTP	YES	YES	YES	**WW39: 15 july 2024**	45	45
WW40	Entrance WWTP	YES	YES	YES	**WW40: 22 july 2024**	33	33
WW41	Entrance WWTP	YES	YES	YES	**WW41: 29 july 2024**	38	35
WW42	Entrance WWTP	YES	YES	YES	**WW42: 05 august 2024**	45	44
WW43	Entrance WWTP	YES	YES	YES	**WW43: 12 august 2024**	45	37
WW44	Entrance WWTP	YES	YES	YES	**WW44: 19 august 2024**	45	37,31
WW45	Entrance WWTP	YES	YES	YES	**WW45: 26 august 2024**	45	36,8
WW46	Entrance WWTP	YES	YES	YES	**WW46: 02 september 2024**	45	37,04
WW47	Entrance WWTP	YES	YES	YES	**WW47: 09 september 2024**	45	36,14

SARS-CoV-2 Lineages

A total of 47 samples were collected using the 3D-printed passive devices. Of these, only 36 samples could be sequenced, as the remaining 11 samples had very late CT values, which prevented successful sequencing. Four sequencing runs were performed using Oxford Nanopore technology (Oxford Nanopore Technologies, 2021). In the first run, 16 samples were sequenced using RNA extracted only from the gauze. In the second run, 4 samples were sequenced while testing RNA extracted from a combination of the swab, gauze, and membrane. The third sequencing run involved 7 samples, again using RNA from the swab, gauze, and membrane combination, but with different primers. Finally, the fourth sequencing run included 9 samples processed under the same conditions as the third. In all sequencing runs, the amplicons generated in the multiplex PCR were diluted at a 1:4 ratio before sequencing.

Figure 11 shows a predominance of the BA.2.86 lineage, which belongs to the Omicron variant of SARS-CoV-2. This lineage was detected in 19 samples from October 2023 to February 2024, consistently showing a relative abundance of over 0.9, indicating reliable results. This lineage reappeared briefly in the week of April 21, 2024, with a relative abundance of 1, and again on August 12, 2024, with a relative abundance above 0.5. The sequencing data were further analyzed using the NEXTCLADE bioinformatics tool. Two samples did not meet the minimum coverage required for analysis and needed to be recognized by the program. However, 14 samples achieved coverage above 50%, confirming the predominance of the JN.1.4.2 PANGO lineage (alias of BA.2.86) during most study weeks. This agreement between data generated by FREYJA and NEXTCLADE indicates reliable lineage assignment.

Between March and July 2024, except for the week of April 21, there was a predominance of the JD.1 lineage and its derivatives, with relative abundances consistently above 0.9. Notably, in the week of June 24, 2024, the XBB.1.23 lineage appeared with a relative abundance greater than 0.5. When comparing these results with NEXTCLADE, some discrepancies were observed; NEXTCLADE assigned samples to the XBB.1.5 and B.1 lineages. However, these assignments were based on coverage percentages below 9%, suggesting that these results may need to be more reliable. It should be noted no insertions or deletions were observed in the sequenced samples, indicating that only the most conserved regions of the virus were sequenced.

From July 15 to September 9, 2024, the XDK lineage and its derivatives were detected in most samples, with relative abundances above 0.9, except for the weeks of July 22 and September 2, when the XDV lineage was also present with relative abundances more significant than 0.5. On August 12, the BA.2.86 lineage reappeared with a relative abundance over 0.7. In comparison, NEXTCLADE assigned most samples to recombinant lineages, with some attributed to the BA.2 lineage. Only one sample, collected on July 15, 2024, was assigned the JN.1 lineage. Coverage percentages for these samples were below 25%, indicating that the NEXTCLADE results may be less reliable than those from FREYJA, where higher relative abundances were observed.

Throughout the study, other lineages were detected in low relative abundances but did not persist over time. These lineages disappeared from one week to the next and were unrelated, which is surprising. This rapid lineage turnover suggests either a shift in circulating variants or changes in primer recognition sites may have affected the results by targeting more conserved viral regions, leading to the detection of ancestral lineages.

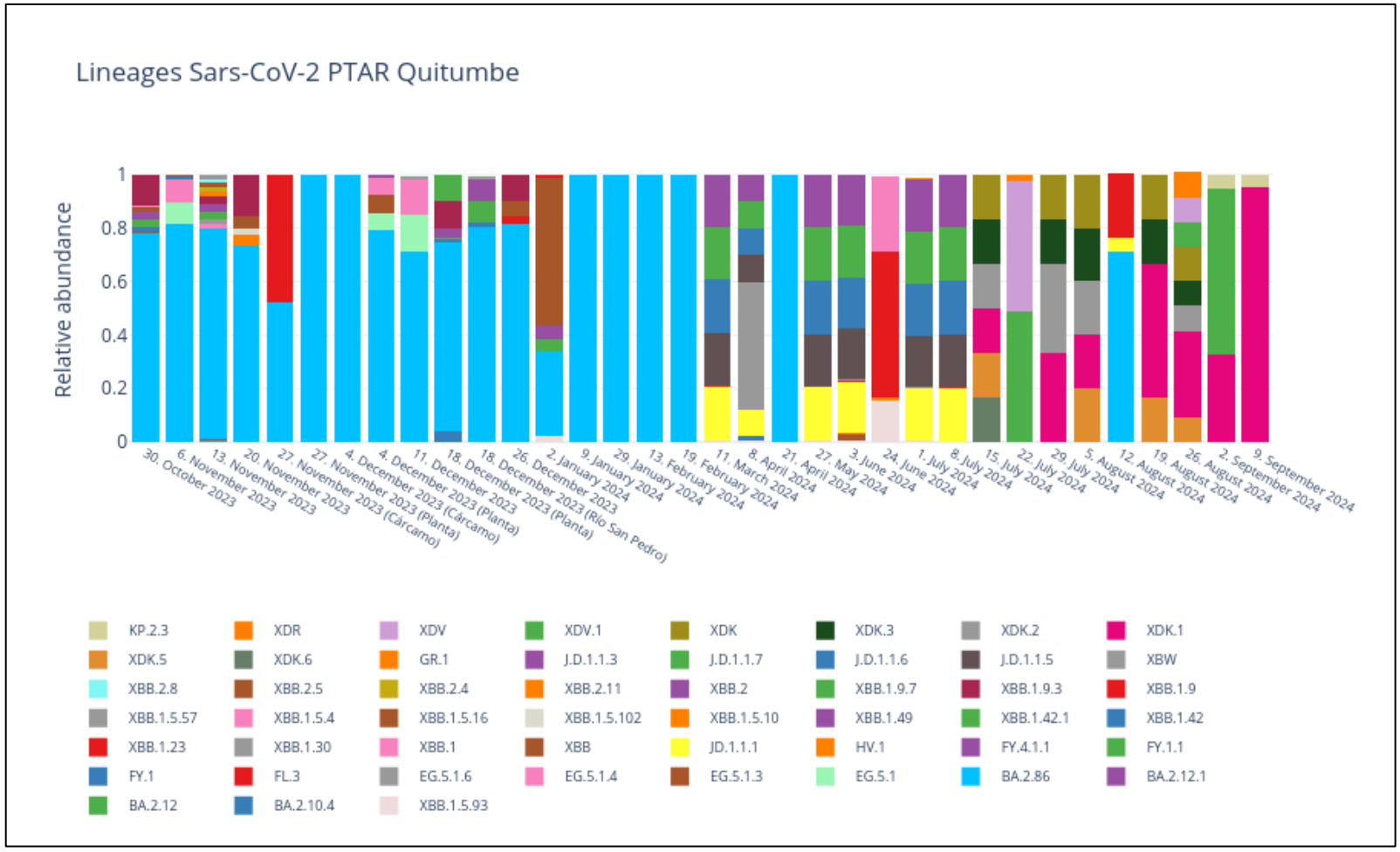

Figure 11 SARS-CoV-2 lineages from PTAR Quitumbe assigned by FREYJA

Comparison of SARS-CoV-2 Lineages Detected in Wastewater and Clinical Samples

Wastewater samples were collected weekly from the Quitumbe Wastewater Treatment Plant between October 2023 and September 2024. These samples were concentrated using the 2X PEG method and processed for RNA extraction using the Zymobiomics Quick-RNA Viral Kit. In contrast, clinical sample data were obtained from the GISAID and CoVariants platforms, where other researchers had uploaded their results, although the exact methods used by these researchers for sample collection and processing are not specified, it is knew that the clinical samples were collected via nasal swabs, followed by RNA extraction and sequencing. The specific hospitals, laboratories, and kits involved in these processes are not known. However, the samples represent various countries, including Argentina, Uruguay, Chile, Paraguay, Bolivia, Brazil, Peru, Ecuador, Colombia, Suriname, Guyana, Venezuela, Aruba, and Trinidad and Tobago. Additionally, the comparison of lineages between wastewater and clinical data showed a strong correlation with the circulating variants reported on GISAID during the same period, further validating the effectiveness of wastewater-based surveillance.. By comparing **Figure 13** and **Figure 14**, it is evident that the XBB.1.5 lineage was detected in wastewater samples from October 2023 to January 2, 2024, except for November 27 and December 4, 2023. This lineage was detected in clinical samples from October 23, 2023, to January 8, 2024. The EG.5.1 lineage was found in wastewater samples from October to December 11, 2023, and clinical samples from October to February 2024. The XBB.1.9 lineage was detected in wastewater from October to November 2023, reappeared in December 2023, and then disappeared, while it was found in clinical samples from November 2023 to January 2024. The HK.3 lineage was not detected in wastewater, probably due to the presence of a host cell in the wastewater samples, due to

its relatively low frequency in clinical samples, where it was detected from 23 November to 6 December 2023 for a short period.

The XBB.1.16 lineage was present in clinical samples from October 2023 to January 2024 but was not detected in wastewater due to its low frequency. The XBB.1.5.70 lineage appeared in clinical samples with a frequency of less than 6% from October 2023 to January 2024, while several low-frequency lineages, XBB.1.5.102, XBB.1.5.16 and XBB.1.5.57, were also detected in wastewater.

The XBB.2.3 lineage was found in clinical samples from October 2023 to January 2024 with a frequency of less than 6%. However, XBB.2 and its sublineages were sporadically detected in wastewater samples from October 2023 to January 2024, although XBB.2.3 was not detected.

The PANGO lineage JN.1 alias BA.2.86 was present in clinical samples from 29 November 2023 to August 2024, while it was detected in wastewater from October to February and again in April when sequencing both POOL A and the combination of the two POOLS. After April, this lineage was no longer detected. The JN.1.11.1 lineage, observed in clinical samples from March to August 2024 with an approximate frequency of 28%, was not detected in wastewater due to primer switching. However, JN.1.4 was detected in April by sequencing only POOL B. The BA.2 lineage was identified in clinical samples from December 2023 to August 2024 at very low frequencies but was not detected in wastewater samples. Instead, its sublineages, such as BA.2.86, BA.2.10.4, BA.2.12, and BA.2.12.1, were detected from October 2023 to February 2024 and March 2024. The KP.3 lineage, which was detectable in clinical samples from June to August 2024, was not observed in any wastewater samples, probably because, as mentioned above, only the conserved part of the virus was being sequenced in samples from March to August, resulting in the assignment of a more primitive lineage.

The recombinant lineages XDK and XDV were detected in wastewater from July 15, 2024, to September 2024. However, these lineages have yet to be observed in clinical samples. The KP.2.3 lineage was identified in wastewater only during September, while a closely related lineage, KP.3, has been detected in clinical samples since June 2024 in low quantities. Notably, the prevalence of KP.3 is increasing over time.

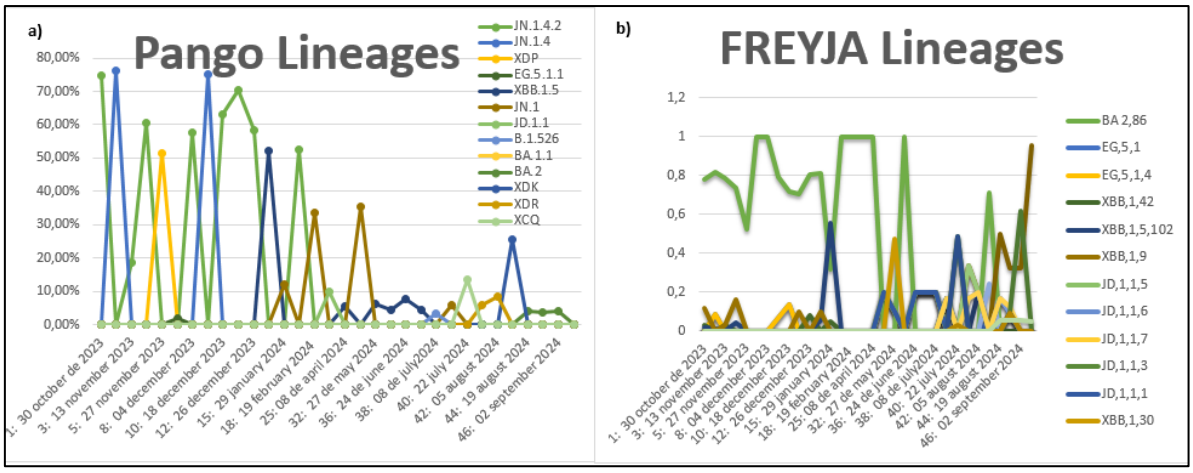

Figure 12 SARS-CoV-2 Lineages Detected in Wastewater Samples from October to September

Panel a shows the lineages identified by the NEXTCLADE tool, corresponding to PANGO lineages, while panel b displays the lineages assigned by the FREYJA tool.

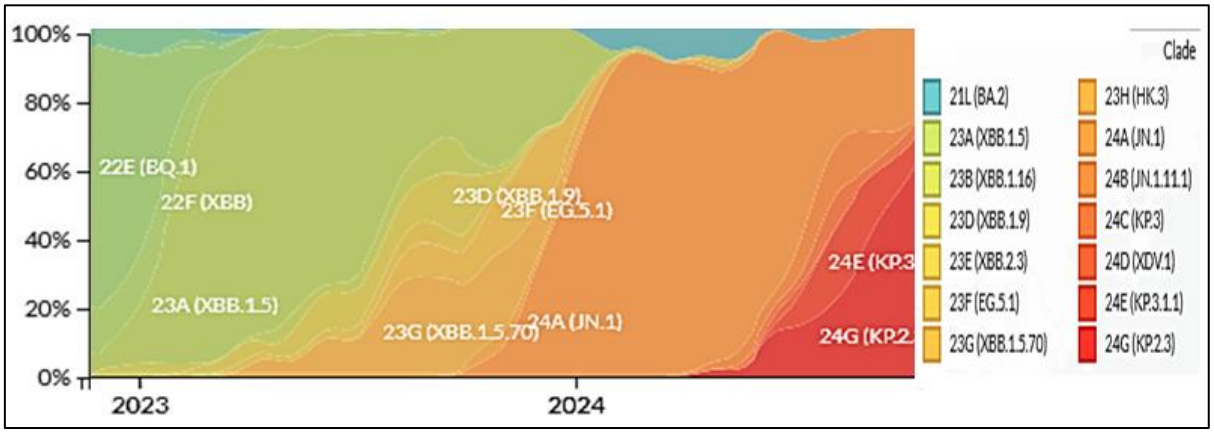

Figure 13 Global SARS-CoV-2 lineages obtained from the GISAID database (GISAID, 2024).

Figure 14 presents data from CoVariants (GISAID, 2024), showing that the Omicron variant has been circulating in Ecuador since October 2023. The identified clades, spanning from the start of the study until January 1, 2024, include 23F, which

corresponds to the PANGO EG.5.1 lineage. This lineage was detected in wastewater between October and November 13, 2023, and then reappeared on December 4 and 11, 2023, before disappearing. The data suggests that the lineage's absence in January is likely due to a reduced viral load, coinciding with its declining presence in clinical samples.

Clade 23H, associated with the HK.3 lineage, was detected at very low levels in clinical samples between October 2023 and January 1, 2024, which may explain why it was not identified in wastewater. However, a recombinant lineage was identified in clinical samples during October 2023; it was at a very low frequency. Conversely, several recombinant lineages were consistently found in wastewater, particularly from October to November 2023, and more prominently from June 24, 2024, until the study's end.

Clade 24B, linked to the JN.1.11.1 lineage, was detected in clinical samples between June and July 15, 2024, but not in wastewater, likely due to its low frequency. Between January and July 15, 2024, clade 24A, associated with the JN.1 lineage, was frequently identified in clinical samples; in wastewater, the JN.1 lineage was detectable from October 2023 to February 2024, and again on April 21 and August 12, 2024, though later detection was limited. Clade 24C, corresponding to the KP.2.3 lineage, has not been detected in Ecuadorian clinical samples but was found in wastewater on September 2 and 9, 2024, at very low relative abundance.

As observed globally, the JN.1 lineage was first detected in wastewater, preceding its appearance in clinical samples by several months. The same phenomenon occurred in our study in Ecuador. The data suggests that wastewater surveillance offers valuable early insights, especially when lineages are circulating at high frequencies and with appropriate primer selection. Furthermore, the reliability of wastewater monitoring is confirmed by the alignment of results with those from clinical samples. However, the sensitivity of the

process to primer changes remains a critical factor in ensuring accurate lineage assignment.

It is also crucial to note that the data on circulating lineages in Ecuador, as shown in **Figure 14**, is only current up to July 15, 2024. Additionally, it remains to be seen whether all researchers involved in genomic surveillance have uploaded their data, which limits the ability to compare findings from this research to the broader dataset fully. Consequently, only a limited number of lineages have been detected compared to those circulating globally.

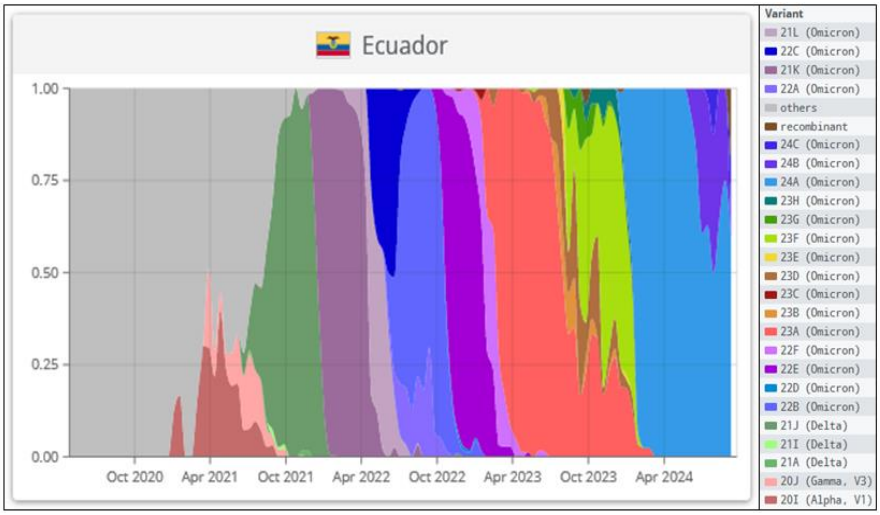

Figure 14 SARS-CoV-2 Strains Circulating in Ecuador Based on CoVariants Data (GISAID, 2024).

Cryptic Wastewater Lineages

According to previous studies (Shafer et al., 2024), epidemiological surveillance of SARS-CoV-2 in wastewater often detects variants that differ from those found in clinical samples. These lineages circulate covertly within the population and are not easily identified by standard surveillance methods due to their low concentration or presence in areas with less genomic monitoring.

These lineages exhibit specific mutations that are uncommon in the most prevalent variants. **Table 2** presents the mutations most frequently identified in wastewater samples, based on studies (Smyth et al., 2022), focusing on the Spike gene. However, as mentioned earlier, the RNA concentration of these cryptic lineages may need to be higher to be detected.

It is also hypothesized that these lineages could originate from a new animal reservoir, such as mammals or rodents that frequently pass through sewers. Detecting these lineages is crucial, as they can provide valuable insights into viral evolution and transmission in specific communities and may help identify new variants of concern before they become widespread.

In this study, cryptic lineages were not detected, which may be attributed to the low concentration of RNA or the possibility that these lineages have yet to emerge in this specific population.

Table 4. Specific mutations in the Spike gene of cryptic Sars CoV-2 lineages (Smyth et al., 2022).

CRYPTIC LINEAGES OF WASTEWATER	E484A/F486P/S494P/Q498Y/H519N/F572N/ Q493K/K440E/F490Y/ Y449R/A346T/V445P/L452Q/T453F/N460L/V483A/F486P /F486V/ K417T/K444T/F590Y/Q498H/N439K/ K444N/Y449R/L452 R/ N460K/S477N/Δ484/F486V/S494T/G496V/ Q498Y/N501T/G504D/Y505H/H519Q

DISCUSSION

Monitoring SARS-CoV-2 remains crucial; one of the key reasons is the early detection of new variants, as an RNA virus mutates rapidly, making ongoing surveillance critical for tracking the virus's evolution, which is essential for public health. In many Latin American countries, clinical surveillance of the virus has become limited, making wastewater monitoring a more feasible and practical alternative to nasal swab testing.

Kitajima et al., 2020, highlight that wastewater surveillance serves as an early warning system for emerging variants of concern, helping to prevent future epidemic waves and allowing for timely updates to vaccines and treatments.

Furthermore, Hodcroft et al., 2021, it underscores the importance of monitoring new mutations to ensure that control measures remain effective, especially as variants with potential immune escape characteristics may arise, Ahmed et al., 2020, it also emphasizes that wastewater surveillance is a valuable complement to clinical testing, as it enables the detection of SARS-CoV-2 in populations that may not undergo regular testing, this noninvasive method provides a more comprehensive understanding of viral circulation, even among asymptomatic individuals or in areas with limited clinical testing resources.

Detection of SARS-CoV-2 in wastewater protocol standardization

Unlike most SARS-CoV-2 epidemiological surveillance studies in wastewater, this study employed a 3D-printed torpedo-type device for sample collection. As described earlier, this device incorporates a membrane, gauze, and swab, distinguishing it from other investigations that collect wastewater directly for subsequent virus enrichment. According to Schang et al., 2021, this 3D-printed device is more affordable and easier to deploy compared to traditional methods. Consistent with Schang's findings, the device in this study proved accessible for sampling from sewers and rivers, and the highest virus concentration was obtained when left in place for 24 continuous hours.

Given that this is a novel sampling method; various virus enrichment protocols were combined. Similar to Woods & Jones, 2021, enhanced viral enrichment was achieved by applying pre-prepared, autoclaved 2X PEG. Additionally, Schang et al., 2021, enrichment process, which involves placing the membrane, swab, and gauze in dPBS with Tween 20, shaking, and incubating for 24 hours, was followed. Our findings

corroborated Schang's protocol, as the best results were obtained when combining these methods after testing different incubation periods and enrichment techniques.

In terms of RNA extraction, O'Brien et al., 2021, compared different kits and identified the Zymo Quick-RNA Viral Kit as the most efficient, particularly for time-saving and RNA preservation. This aligns with our research, where the Zymo kit produced the best results compared to other kits.

For epidemiological surveillance of SARS-CoV-2 in wastewater, Kazenelson et al., 2023, the virus was detected using qPCR, confirming the presence of the nucleocapsid's N1 and N2 genes. Similarly, in this study, both N1 and N2 genes were detected in 20 out of 47 samples, while in the remaining samples, only one of the two genes was detected.

To generate amplicons for multiplex PCR, Barnes et al., 2023, used the NEBNext® Varskip Short ARTIC SARS-CoV-2 primer set. Over a two-year period, their study demonstrated that new lineages were first detected in wastewater samples before appearing in clinical samples. Likewise, in this study, the BA.2.86 lineage was detected earlier in wastewater samples using these primers. However, when the primers were changed, the correct detection of lineages diminished, likely due to improper adherence of the primers to the virus, complicating accurate lineage assignment. Parra-Guardado et al., 2022, noted that designing primers for specific lineages allows for more precise detection. Though the primers were not custom designed in this study due to time constraints and standardization requirements, the literature suggests that this could introduce bias, limiting detection to specific lineages and potentially missing cryptic lineages.

In terms of sequencing, Tyson et al., 2020, they demonstrated that Oxford Nanopore Technologies optimized both the library workflow and costs, making it more accessible for genomic surveillance than other sequencing platforms. Crits-Christoph et al., 2021, further highlighted the rapid identification of new mutations in wastewater using Nanopore, a feature echoed in this study where real-time detection of SARS-CoV-2 lineages was achieved, saving both time and resources.

qPCR quantifications

Ct values vary depending on viral load, environmental factors, and the detection method used. Nauta et al., 2023, showed that variations in viral concentration can influence Ct values, which are affected by factors such as sampling frequency, RNA degradation, and patterns of viral shedding. In contrast, studies such as Ahmed et al., 2020, show that late Ct values are common when viral concentrations in wastewater are low, particularly in areas with reduced transmission, which may indicate a decreasing or residual viral presence, but they are still valuable for identifying the virus.

O'Brien et al., 2021, also highlighted that samples with late Ct values, which indicate low concentrations of viral RNA, can still be critical for detecting the virus and monitoring outbreaks, even when detection becomes more difficult. Comparing our results to these studies allows us to assess how late Ct values provide insight into viral trends. This offers a basis for understanding viral circulation, especially when direct clinical testing is limited.

As shown in **Table 1**, approximately 15 of the 47 samples processed for the N1 and N2 genes had late Ct values (>38). Most studies reviewed only samples sequenced with Ct values less than 35, but in this study, samples with higher Ct values were sequenced. Furthermore, the pattern of early detection for the N1 gene followed by late

detection and a similar trend for the N2 gene (detected late and then early) may suggest viral evolutionary dynamics, with the virus adapting to persist over time. Factors such as sampling methods, viral concentration techniques, and community prevalence of the virus are likely to play an important role.

The literature confirms that SARS-CoV-2 can still be detected in wastewater even when viral concentrations are low, allowing for the tracking of new outbreaks. In addition, factors such as polymerase selection, primers, and amplification efficiency must be considered. As Tiwari et al., 2023, demonstrated in their study, detection is not limited to the N1 and N2 genes; they also amplified the RdRP and E genes, as recommended by the CDC, but in this study only detection N1 and N2 genes.

SARS-CoV-2 Lineages

Yousif et al., 2023, highlight that whole genome sequencing can be applied to wastewater to detect and characterize SARS-CoV-2 variants, enabling the recovery of full viral genomes, observation of virus dynamics by comparing clinical samples with wastewater samples, and identifying new mutations first in wastewater and later in clinical samples. They also emphasize the importance of applying this methodology in low- and middle-income countries or regions with limited access to clinical testing.

In our study, we can corroborate Yousif et al.'s findings, as shown in Figure 10, lineage BA.2.86, also known as Pirola, was detected early in our investigation—starting in October 2023, however, clinical samples first identified it on November 29, 2023, nearly four weeks later. Similarly, in Yousif et al.'s study, high RNA concentrations in wastewater or high prevalence in the population allowed for earlier detection of a lineage. However, lineages at low prevalence in the population had minimal RNA concentrations, making them undetectable in wastewater even with enrichment processes. As in Yousif

et al.'s research, there were periods in our study where specific lineages could not be sequenced due to insufficient reads for lineage assignment.

In this study, we successfully assigned lineages to all sequenced samples by selecting them based on their Ct values obtained through qPCR, a critical aspect to highlight is the use of primers for example from March 2024 to the first week of July 2024, different primers were used, resulting in the detection of primitive lineages—those observed early in the pandemic. However, in samples from the second week of July to September 2024, when we switched back to the original primers, we observed recombinant lineages and clinical samples. BA.2.86 continued to dominate, but an interesting finding was the detection of KP.2.3 (a sublineage of JN.1) in wastewater starting on September 2, 2024, and continuing through the end of the study. Meanwhile, clinical samples detected another JN.1 sublineage, KP.3, with low prevalence since June 2024, which increased over time. This demonstrates the potential of wastewater surveillance to detect SARS-CoV-2 lineages and monitor viral evolution.

In a study by Kaku et al., 2024, newer sublineages such as KP.2, KP.3, and KP.2.3, along with LB.1, emerged later in 2024 with additional spike protein mutations. These variants display evolutionary traits, particularly in immune evasion. The combined use of clinical and wastewater surveillance enhances our understanding of viral dynamics, for example, KP.3 was only detected in clinical samples, but KP.2.3 was consistently found in wastewater samples, albeit at low relative abundance, indicating potential viral evolution and rapid acquisition of mutations.

Furthermore, Kaku et al., 2024, highlight the reproduction number (R_e), which measures how quickly a virus can spread. KP.3 and KP.2.3 had a significantly higher R_e than their parental lineage, JN.1. In terms of infectivity, KP.3 was less infectious than JN.1, while LB.1 and KP.2.3 exhibited similar infectivity to JN.1. Kaku et al., 2024, also

notes that KP.2.3 and LB.1 were more resistant to neutralization compared to JN.1, suggesting that these variants have greater potential to escape immunity, even in individuals previously infected with variants like XBB or those who have been vaccinated. These data suggest that KP.2.3, LB.1, and KP.3 are likely to spread globally and contribute to future outbreaks. Given the virus's rapid evolution and its ability to acquire new mutations, as demonstrated in this study, it is crucial to maintain genomic surveillance in wastewater to track emerging variants that may challenge current vaccines and treatments.

It’s important to note that in Ecuador, only one molecular detection study for SARS-CoV-2 in wastewater has been conducted which utilized RT-qPCR to detect the presence or absence of the virus. Delgado- Salgado et al., 2023, however, this method does not allow for analyzing viral dynamics or lineage circulation. Despite the proven utility of genomic surveillance in wastewater as a tool for detecting viral circulation and anticipating outbreaks, continuous implementation in many countries, including those in Latin America, has been limited. This can be attributed to factors such as a lack of technological infrastructure, insufficient funding, and the absence of public policies prioritizing this type of surveillance outside of emergencies.

Cryptic lineages

In this study, cryptic SARS-CoV-2 lineages were not detected in wastewater samples, which can be attributed to several technical and epidemiological factors. Shafer et al., 2024, conducted a longitudinal analysis of urban wastewater and identified an Omicron-like variant that did not correspond to clinical sequences recorded during the same period. This discovery highlights the ability of wastewater surveillance to capture emerging viral variants that may go unnoticed in clinical testing, particularly when cryptic lineages are circulating undetected within the community, however, the presence of

cryptic lineages has been observed in specific settings, often in areas with high infection prevalence or in studies that employ consistent, long-term genomic surveillance. However, in the context of this work, the methods used—while robust for detecting predominant variants like BA.2.86 and KP.2.3—may not have been sufficient to identify cryptic lineages.

Additionally, studies in the United States have demonstrated the potential of wastewater surveillance to detect cryptic lineages when regular and prolonged genomic monitoring is in place (Smyth et al., 2022). These studies also point out that low-prevalence lineages, or those in the early stages of emergence, may not be detected due to the minimal viral load in the samples. In the case of this study, the Ct values used to select sequenced samples likely prioritized those with higher viral loads, which may have excluded low-abundance lineages.

The absence of cryptic lineages can also be interpreted as an indicator that, during the study period, the predominant lineages were well represented in the wastewater samples. This is consistent with the prevalence of lineages like BA.2.86 in both clinical and wastewater samples, suggesting lower viral diversity and effective epidemiological control of the virus in the population during the study period. The ability to detect emerging lineages, such as KP.2.3, and track their evolution underscores the importance of genomic surveillance as a tool for timely detection of variants, even when cryptic lineages are not identified.

CONCLUSION AND LIMITATIONS

This study introduced a significant innovation in sample collection for SARS-CoV-2 surveillance in wastewater through the use of a 3D-printed torpedo-type device. This device provides distinct advantages over conventional sampling methods, as it facilitates the collection of viral material via swabs, membranes, and gauze rather than

directly from wastewater; this approach improves virus recovery by reducing RNA dilution, which is a common challenge in traditional methods. Additionally, it allows for more targeted viral particle collection, enhancing detection sensitivity, particularly in areas without infrastructure for continuous wastewater monitoring.

The novel method proved effective for the genomic surveillance of predominant SARS-CoV-2 lineages, such as BA.2.86 and KP.2.3, underscoring its value as an epidemiological tool to track viral evolution in a population. Our results align with previous studies that demonstrate the ability of wastewater surveillance to detect emerging variants earlier than in clinical settings, providing a critical advantage for public health interventions.

However, this study did not detect cryptic lineages, which may be attributed to technical factors such as the prioritization of samples with higher viral loads (Ct values). Prolonged surveillance, as demonstrated by Shafer et al., can identify cryptic lineages, but the relatively short duration of this study, combined with the focus on high viral load samples, may have limited the detection of lower-frequency lineages. As Smyth et al. have suggested, cryptic lineages with lower viral loads may require more consistent and long-term monitoring to be effectively detected.

One limitation encountered was the change in primers used between March and July 2024, which impacted the detection of circulating lineages and led to the identification of primitive lineages during that period, this underscores the importance of maintaining consistent, updated methodologies in genomic sequencing to ensure accurate lineage representation over time.

Additional limitations should be considered when replicating this methodology in other contexts. The need to process samples immediately after device deployment to

prevent RNA degradation is critical. Furthermore, the short monitoring period and the previously primer change affected the accuracy of lineage assignments, limiting the detection of some variants at specific times.

REFERENCES

Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J. W., Choi, P. M., Kitajima, M., Simpson, S. L., Li, J., Tscharke, B., Verhagen, R., Smith, W. J. M., Zaugg, J., Dierens, L., Hugenholtz, 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. *Science of the Total Environment*, *728*, 138764. https://doi.org/10.1016/j.scitotenv.2020.138764

Aksamentov, I., Roemer, C., Hodcroft, E., & Neher, R. (2021). Nextclade: clade assignment, mutation calling and quality control for viral genomes. *Journal of Open Source Software*, *6*(67), 3773. https://doi.org/10.21105/joss.03773

Barnes, K. G., Levy, J. I., Gauld, J., Rigby, J., Kanjerwa, O., Uzzell, C. B., Chilupsya, C., Anscombe, C., Tomkins-Tinch, C., Mbeti, O., Cairns, E., Thole, H., McSweeney, S., Chibwana, M. G., Ashton, P. M., Jere, K. C., Meschke, J. S., Diggle, P., Cornick, J., … Feasey, N. (2023). Utilizing river and wastewater as a SARS-CoV-2 surveillance tool in settings with limited formal sewage systems. *Nature Communications*, *14*(1), 1–11. https://doi.org/10.1038/s41467-023-43047-y

Callaway, E. (2022). These scientists traced a new coronavirus lineage to one office - through sewage. *Nature*, *609*(7929), 883–885. https://doi.org/10.1038/d41586-022-02996-y

Campbell, F., Archer, B., Laurenson-Schafer, H., Jinnai, Y., Konings, F., Batra, N., Pavlin, B., Vandemaele, K., Van Kerkhove, M. D., Jombart, T., Morgan, O., & De Waroux, O. L. P. (2021). Increased transmissibility and global spread of SARSCoV-2 variants of concern as at June 2021. *Eurosurveillance*, *26*(24), 1–6.

https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100509

Candido, D. S., Claro, I. M., de Jesus, J. G., Souza, W. M., Moreira, F. R. R., Dellicour, S., Mellan, T. A., du Plessis, L., Pereira, R. H. M., Sales, F. C. S., Manuli, E. R., Thézé, J., Almeida, L., Menezes, M. T., Voloch, C. M., Fumagalli, M. J., Coletti, T. M., da Silva, C. A. M., Ramundo, M. S., … Faria, N. R. (2020). Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*, *369*(6508), 1255–1260. https://doi.org/10.1126/SCIENCE.ABD2161

CDC. (2020). *National Wastewater Surveillance System (NWSS)*. 1–4. (Accessed 09/08/2023)

Centers for Disease Control and Prevention. (2024). *Wastewater COVID-19 National and Regional Trends*. 1–5. https://www.cdc.gov/nwss/rv/COVID19-nationaltrend.html

Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K. W., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M. L., Mulders, D. G. J. C., Haagmans, B. L., Van Der Veer, B., Van Den Brink, S., Wijsman, L., Goderski, G., Romette, J. L., Ellis, J., Zambon, M., … Drosten, C. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, *25*(3). https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045

Cotten, M., Jaimes, J. A., & Whittaker, R. (n.d.). *Spike Protein Cleavage-Activation in the Context of the*.

Crits-Christoph, A., Kantor, R. S., Olm, M. R., Whitney, O. N., Al-Shayeb, B., Lou, Y. C., Flamholz, A., Kennedy, L. C., Greenwald, H., Hinkle, A., Hetzel, J., Spitzer, S., Koble, J., Tan, A., Hyde, F., Schroth, G., Kuersten, S., Banfield, J. F., & Nelson, K. L. (2021). Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. *MBio*, *12*(1), 1–9. https://doi.org/10.1128/mBio.02703-20

Delgado- Salgado, P., Champutiz-Ortiz, E., Muñoz-Vera, C., & Asimbaya-Alvarado, D. (2023). Detección molecular de SARS-CoV-2 en aguas residuales como componente de vigilancia ambiental de COVID-19. - Molecular detection of SARS-CoV-2 in wastewater as a component of environmental surveillance of the COVID-19. *Revista de La Facultad de Ciencias Médicas (Quito)*, *48*(2), 9–15. https://doi.org/10.29166/rfcmq.v48i2.5778

GISAID. (2022). Enabling rapid and open access to epidemic and pandemic virus data. *Gisaid*, 8–9. https://www.gisaid.org/about-us/mission/

Greaney, A. J., Loes, A. N., Crawford, K. H. D., Starr, T. N., Malone, K. D., Chu, H. Y., & Bloom, J. D. (2021). Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host and Microbe*, *29*(3), 463-476.e6. https://doi.org/10.1016/j.chom.2021.02.003

Gregory, D. A., Rushford, C., Hunter, T., Lin, C. H., Darby, C., Niehues, N., Semkiw, E., Reynolds, M., Wenzel, J., & Johnson, M. C. (2023). Continued selection on cryptic SARS-CoV-2 observed in Missouri wastewater. *PLoS Pathogens*, *19*(12 December), 1–8. https://doi.org/10.1371/journal.ppat.1011688

Gregory, D. A., Trujillo, M., Rushford, C., Flury, A., Kannoly, S., San, K. M., Lyfoung, D. T., Wiseman, R. W., Bromert, K., Zhou, M. Y., Kesler, E., Bivens, N. J., Hoskins, J., Lin, C. H., O'Connor, D. H., Wieberg, C., Wenzel, J., Kantor, R. S., Dennehy, J. J., & Johnson, M. C. (2022). Genetic diversity and evolutionary convergence of cryptic SARS- CoV-2 lineages detected via wastewater sequencing. *PLoS Pathogens*, *18*(10), 1–25. https://doi.org/10.1371/journal.ppat.1010636

Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., Sagulenko,

P., Bedford, T., & Neher, R. A. (2018). Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*, *34*(23), 4121–4123. https://doi.org/10.1093/bioinformatics/bty407

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. *Nature Reviews Microbiology*, *19*(7), 409–424. https://doi.org/10.1038/s41579-021-00573-0

Hodcroft, E. B., Zuber, M., Nadeau, S., Vaughan, T. G., Crawford, K. H. D., Althaus, C. L., Reichmuth, M. L., Bowen, J. E., Walls, A. C., Corti, D., Bloom, J. D., Veesler, D., Mateo, D., Hernando, A., Comas, I., González-Candelas, F., González-Candelas, F., Goig, G. A., Chiner-Oms, Á., … Neher, R. A. (2021). Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature*, *595*(7869), 707–712. https://doi.org/10.1038/s41586-021-03677-y

Kaku, Y., Yo, M. S., Tolentino, J. E., Uriu, K., Okumura, K., Ito, J., & Sato, K. (2024). Virological characteristics of the SARS-CoV-2 KP.3, LB.1, and KP.2.3 variants. *The Lancet Infectious Diseases*, *24*(8), e482–e483. https://doi.org/10.1016/S1473-3099(24)00415-8

Karthikeyan, S., Levy, J. I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., Tribelhorn, C. E., Tsai, R., Aigner, S., Sathe, S., Moshiri, N., Henson, B., Mark, A. M., Hakim, A., Baer, N. A., Barber, T., … Knight, R. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature*, *609*(7925), 101–108. https://doi.org/10.1038/s41586-022-05049-6

Kazenelson, J., Jefferson, T., Rhodes, R. G., Cahoon, L. B., & Frampton, A. R. (2023).

Detection of SARS-CoV-2 RNA in wastewater from an enclosed college campus serves as an early warning surveillance system. *PLoS ONE*, *18*(7 July), 1–15. https://doi.org/10.1371/journal.pone.0288808

Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C. P., Hamilton, K. A., Haramoto, E., & Rose, J. B. (2020). SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Science of the Total Environment*, *739*, 139076. https://doi.org/10.1016/j.scitotenv.2020.139076

Kitamura, K., Sadamasu, K., Muramatsu, M., & Yoshida, H. (2021). Efficient detection of SARS-CoV-2 RNA in the solid fraction of wastewater. *Science of the Total Environment*, *763*, 144587. https://doi.org/10.1016/j.scitotenv.2020.144587

Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E. E., Bhattacharya, T., Foley, B., Hastie, K. M., Parker, M. D., Partridge, D. G., Evans, C. M., Freeman, T. M., de Silva, T. I., Angyal, A., Brown, R. L., Carrilero, L., … Montefiori, D. C. (2020). Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*, *182*(4), 812-827.e19. https://doi.org/10.1016/j.cell.2020.06.043

Long, Q. X., Liu, B. Z., Deng, H. J., Wu, G. C., Deng, K., Chen, Y. K., Liao, P., Qiu, J. F., Lin, Y., Cai, X. F., Wang, D. Q., Hu, Y., Ren, J. H., Tang, N., Xu, Y. Y., Yu, L. H., Mo, Z., Gong, F., Zhang, X. L., … Huang, A. L. (2020). Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nature Medicine*, *26*(6), 845–848. https://doi.org/10.1038/s41591-020-0897-1

Motozono, C., Toyoda, M., Zahradnik, J., Saito, A., Nasser, H., Tan, T. S., Ngare, I., Kimura, I., Uriu, K., Kosugi, Y., Yue, Y., Shimizu, R., Ito, J., Torii, S., Yonekawa, A., Shimono, N., Nagasaki, Y., Minami, R., Toya, T., … Sato, K. (2021). SARS-

CoV-2 spike L452R variant evades cellular immunity and increases infectivity. *Cell Host and Microbe*, *29*(7), 1124-1136.e11. https://doi.org/10.1016/j.chom.2021.06.006

Nauta, M., McManus, O., Træholt Franck, K., Lindberg Marving, E., Dam Rasmussen, L., Raith Richter, S., & Ethelberg, S. (2023). Early detection of local SARS-CoV-2 outbreaks by wastewater surveillance: A feasibility study. *Epidemiology and Infection*, *151*. https://doi.org/10.1017/S0950268823000146

Negi, S. S., Schein, C. H., & Braun, W. (2022). Regional and temporal coordinated mutation patterns in SARS-CoV-2 spike protein revealed by a clustering and network analysis. *Scientific Reports*, *12*(1), 1–10. https://doi.org/10.1038/s41598-022-04950-4

O'Brien, M., Rundell, Z. C., Nemec, M. D., Langan, L. M., Back, J. A., & Lugo, J. N. (2021). A comparison of four commercially available RNA extraction kits for wastewater surveillance of SARS-CoV-2 in a college population. *Science of the Total Environment*, *801*, 149595. https://doi.org/10.1016/j.scitotenv.2021.149595

Oude Munnink, B. B., Nieuwenhuijse, D. F., Stein, M., O'Toole, Á., Haverkate, M., Mollers, M., Kamga, S. K., Schapendonk, C., Pronk, M., Lexmond, P., van der Linden, A., Bestebroer, T., Chestakova, I., Overmars, R. J., van Nieuwkoop, S., Molenkamp, R., van der Eijk, A. A., GeurtsvanKessel, C., Vennema, H., … Hoebe, C. (2020). Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nature Medicine*, *26*(9), 1405–1410. https://doi.org/10.1038/s41591-020-0997-y

Parra-Guardado, A. L., Sweeney, C. L., Hayes, E. K., Trueman, B. F., Huang, Y., Jamieson, R. C., Rand, J. L., Gagnon, G. A., & Stoddart, A. K. (2022). Development

of a rapid pre-concentration protocol and a magnetic beads-based RNA extraction method for SARS-CoV-2 detection in raw municipal wastewater. *Environmental Science: Water Research and Technology*, *8*(1), 47–61. https://doi.org/10.1039/d1ew00539a

Peccia, J., Zulli, A., Brackney, D. E., Grubaugh, N. D., Kaplan, E. H., Casanovas-Massana, A., Ko, A. I., Malik, A. A., Wang, D., Wang, M., Warren, J. L., Weinberger, D. M., Arnold, W., & Omer, S. B. (2020). Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nature Biotechnology*, *38*(10), 1164–1167. https://doi.org/10.1038/s41587-020-0684-z

Planas, D., Bruel, T., Grzelak, L., Guivel-Benhassine, F., Staropoli, I., Porrot, F., Planchais, C., Buchrieser, J., Rajah, M. M., Bishop, E., Albert, M., Donati, F., Prot, M., Behillil, S., Enouf, V., Maquart, M., Smati-Lafarge, M., Varon, E., Schortgen, F., … Schwartz, O. (2021). Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nature Medicine*, *27*(5), 917–924. https://doi.org/10.1038/s41591-021-01318-5

Rambaut, A., Holmes, E. C., O'Toole, Á., Hill, V., McCrone, J. T., Ruis, C., du Plessis, L., & Pybus, O. G. (2020). A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology*, *5*(11), 1403–1407. https://doi.org/10.1038/s41564-020-0770-5

Ramos-Montañez, S., Tsui, H. C. T., Wayne, K. J., Morris, J. L., Peters, L. E., Zhang, F., Kazmierczak, K. M., Sham, L. T., & Winkler, M. E. (2008). Polymorphism and regulation of the spxB (pyruvate oxidase) virulence factor gene by a CBS-HotDog domain protein (SpxR) in serotype 2 Streptococcus pneumoniae. *Molecular Microbiology*, *67*(4), 729–746. https://doi.org/10.1111/j.1365-2958.2007.06082.x

Santos, I. de A., Grosche, V. R., Bergamini, F. R. G., Sabino-Silva, R., & Jardim, A. C. G. (2020). Antivirals Against Coronaviruses: Candidate Drugs for SARS-CoV-2 Treatment? *Frontiers in Microbiology*, *11*(August). https://doi.org/10.3389/fmicb.2020.01818

Schang, C., Crosbie, N. D., Nolan, M., Poon, R., Wang, M., Jex, A., John, N., Baker, L., Scales, P., Schmidt, J., Thorley, B. R., Hill, K., Zamyadi, A., Tseng, C. W., Henry, R., Kolotelo, P., Langeveld, J., Schilperoort, R., Shi, B., … McCarthy, D. T. (2021). Passive Sampling of SARS-CoV-2 for Wastewater Surveillance. *Environmental Science and Technology*, *55*(15), 10432–10441. https://doi.org/10.1021/acs.est.1c01530

Schoeman, D., Fielding, B. C., Arias-Reyes, C., Zubieta-DeUrioste, N., Poma-Machicao, L., Aliaga-Raudan, F., Carvajal-Rodriguez, F., Dutschmann, M., Schneider-Gasser, E. M., Zubieta-Calleja, G., Soliz, J., Schneider-Gasser, E. M., Zubieta-Calleja Director High, G., Loeffelholz, M. J., Tang, Y. W., Velavan, T. P., Meyer, C. G., Taylor, D., Lindsay, A. C., … Sheraton, K. (2020). Journal Pre-proof Does the pathogenesis of SAR-CoV-2 virus decrease at high-altitude? Does the pathogenesis of SAR-CoV-2 virus decrease at high-altitude? Corresponding authors. *Cell Research*, *9*(1), 278–280. https://doi.org/10.3390/ijerph17082932

Scohy, A., Anantharajah, A., Bodéus, M., Kabamba-Mukadi, B., Verroken, A., & Rodriguez-Villalobos, H. (2020). Low performance of rapid antigen detection test as frontline testing for COVID-19 diagnosis. *Journal of Clinical Virology*, *129*(May), 104455. https://doi.org/10.1016/j.jcv.2020.104455

Shafer, M. M., Bobholz, M. J., Vuyk, W. C., Gregory, D. A., Roguet, A., Haddock Soto, L. A., Rushford, C., Janssen, K. H., Emmen, I. E., Ries, H. J., Pilch, H. E., Mullen,

P. A., Fahney, R. B., Wei, W., Lambert, M., Wenzel, J., Halfmann, P., Kawaoka, Y., Wilson, N. A., … Johnson, M. C. (2024). Tracing the origin of SARS-CoV-2 omicron-like spike sequences detected in an urban sewershed: a targeted, longitudinal surveillance study of a cryptic wastewater lineage. *The Lancet Microbe*, *5*(4), e335–e344. https://doi.org/10.1016/S2666-5247(23)00372-5

Smyth, D. S., Trujillo, M., Gregory, D. A., Cheung, K., Gao, A., Graham, M., Guan, Y., Guldenpfennig, C., Hoxie, I., Kannoly, S., Kubota, N., Lyddon, T. D., Markman, M., Rushford, C., San, K. M., Sompanya, G., Spagnolo, F., Suarez, R., Teixeiro, E., … Dennehy, J. J. (2022). Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nature Communications*, *13*(1), 1–9. https://doi.org/10.1038/s41467-022-28246-3

Starr, T. N., Greaney, A. J., Hilton, S. K., Ellis, D., Crawford, K. H. D., Dingens, A. S., Navarro, M. J., Bowen, J. E., Tortorici, M. A., Walls, A. C., King, N. P., Veesler, D., & Bloom, J. D. (2020). Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell*, *182*(5), 1295-1310.e20. https://doi.org/10.1016/j.cell.2020.08.012

Surjit, M., & Lal, S. K. (2008). The SARS-CoV nucleocapsid protein: A protein with multifarious activities. *Infection, Genetics and Evolution*, *8*(4), 397–405. https://doi.org/10.1016/j.meegid.2007.07.004

Tegally, H., Moir, M., Everatt, J., Giovanetti, M., Scheepers, C., Wilkinson, E., Subramoney, K., Makatini, Z., Moyo, S., Amoako, D. G., Baxter, C., Althaus, C. L., Anyaneji, U. J., Kekana, D., Viana, R., Giandhari, J., Lessells, R. J., Maponga, T., Maruapula, D., … de Oliveira, T. (2022). Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in South Africa. *Nature Medicine*, *28*(9), 1785–1790.

https://doi.org/10.1038/s41591-022-01911-2

Tiwari, A., Adhikari, S., Zhang, S., Solomon, T. B., Lipponen, A., Islam, M. A., Thakali, O., Sangkham, S., Shaheen, M. N. F., Jiang, G., Haramoto, E., Mazumder, P., Malla, B., Kumar, M., Pitkänen, T., & Sherchan, S. P. (2023). Tracing COVID-19 Trails in Wastewater: A Systematic Review of SARS-CoV-2 Surveillance with Viral Variants. *Water (Switzerland)*, *15*(6), 1–19. https://doi.org/10.3390/w15061018

Tyson, J. R., James, P., Stoddart, D., Sparks, N., Wickenhagen, A., Hall, G., Choi, J. H., Lapointe, H., Kamelian, K., Smith, A. D., Prystajecky, N., Goodfellow, I., Wilson, S. J., Harrigan, R., Snutch, T. P., Loman, N. J., & Quick, J. (2020). Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *BioRxiv : The Preprint Server for Biology*. https://doi.org/10.1101/2020.09.04.283077

Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*, *181*(2), 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058

Woods, J., & Jones, J. L. (2021). *2x peg*. 1–2.

World Health Organization, W. (2021). *Tracking Variants*. *May*, 1–10. https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/COVID-19/COVID-Variants.aspx

Worobey, M., Pekar, J., Larsen, B. B., Nelson, M. I., Hill, V., Joy, J. B., Rambaut, A., Suchard, M. A., Wertheim, J. O., & Lemey, P. (2020). The emergence of SARS-CoV-2 in Europe and North America. *Science*, *370*(6516), 564–570. https://doi.org/10.1126/SCIENCE.ABC8169

Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., Graham, B. S., & McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, *367*(6483), 1260–1263. https://doi.org/10.1126/science.aax0902

Wurtzer, S., Marechal, V., Mouchel, J. M., Maday, Y., Teyssou, R., Richard, E., Almayrac, J. L., & Moulin, L. (2020). Evaluation of lockdown effect on SARS-CoV-2 dynamics through viral genome quantification in waste water, Greater Paris, France, 5 March to 23 April 2020. *Eurosurveillance*, *25*(50). https://doi.org/10.2807/1560-7917.ES.2020.25.50.2000776

Yousif, M., Rachida, S., Taukobong, S., Ndlovu, N., Iwu-Jaja, C., Howard, W., Moonsamy, S., Mhlambi, N., Gwala, S., Levy, J. I., Andersen, K. G., Scheepers, C., von Gottberg, A., Wolter, N., Bhiman, J. N., Amoako, D. G., Ismail, A., Suchard, M., & McCarthy, K. (2023). SARS-CoV-2 genomic surveillance in wastewater as a model for monitoring evolution of endemic viruses. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-41369-5

Zhang, D., Duran, S. S. F., Lim, W. Y. S., Tan, C. K. I., Cheong, W. C. D., Suwardi, A., & Loh, X. J. (2022). SARS-CoV-2 in wastewater: From detection to evaluation. *Materials Today Advances*, *13*, 100211. https://doi.org/10.1016/j.mtadv.2022.100211

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733. https://doi.org/10.1056/nejmoa2001017