All Faculty Scholarship

4-21-2023

Wastewater Surveillance of SARS-CoV-2 Genomic Populations on a Country-wide Scale Through Targeted Sequencing

Florencia Cancela

Natalia Ramos

Davida S. Smyth

et al.

PLOS ONE

Wastewater surveillance of SARS-CoV-2 genomic populations on a country-wide scale through targeted sequencing

Florencia Cancela[1], **Natalia Ramos**[1], **Davida S. Smyth**[2], **Claudia Etchebehere**[3], **Mabel Berois**[1], **Jesica Rodríguez**[4], **Caterina Rufo**[4], **Alicia Alemán**[5], **Liliana Borzacconi**[6], **Julieta López**[7], **Elizabeth González**[7], **Germán Botto**[8], **Starla G. Thornhill**[2], **Santiago Mirazo**[1,9]*, **Mónica Trujillo**[10]*

1 Sección Virología, Instituto de Química Biológica, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay, 2 Department of Life Sciences, Texas A&M University-San Antonio, San Antonio, Texas, United States of America, 3 Departamento de Bioquímica y Genómica Microbiana, Instituto de Investigaciones Biológicas Clemente Estable, Ministerio de Educación y Cultura, Montevideo, Uruguay, 4 Laboratorio de Alimentos y Nutrición, Polo Tecnológico de Pando, Facultad de Química, Universidad de la República, Montevideo, Uruguay, 5 Departamento de Medicina Preventiva, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay, 6 Instituto de Ingeniería Química, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay, 7 Departamento de Ingeniería Ambiental, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay, 8 Departamento de Métodos Cuantitativos, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay, 9 Departamento de Bacteriología y Virología, Instituto de Higiene, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay, 10 Department of Biological Sciences and Geology, Queensborough Community College of The City University of New York, Queens, New York, United States of America

* mtrujillo@qcc.cuny.edu (MT); smirazo@higiene.edu.uy (SM)

Abstract

SARS-CoV-2 surveillance of viral populations in wastewater samples is recognized as a useful tool for monitoring epidemic waves and boosting health preparedness. Next generation sequencing of viral RNA isolated from wastewater is a convenient and cost-effective strategy to understand the molecular epidemiology of SARS-CoV-2 and provide insights on the population dynamics of viral variants at the community level. However, in low- and middle-income countries, isolated groups have performed wastewater monitoring and data has not been extensively shared in the scientific community. Here we report the results of monitoring the co-circulation and abundance of variants of concern (VOCs) of SARS-CoV-2 in Uruguay, a small country in Latin America, between November 2020—July 2021 using wastewater surveillance. RNA isolated from wastewater was characterized by targeted sequencing of the Receptor Binding Domain region within the spike gene. Two computational approaches were used to track the viral variants. The results of the wastewater analysis showed the transition in the overall predominance of viral variants in wastewater from No-VOCs to successive VOCs, in agreement with clinical surveillance from sequencing of nasal swabs. The mutations K417T, E484K and N501Y, that characterize the Gamma VOC, were detected as early as December 2020, several weeks before the first clinical case was reported. Interestingly, a non-synonymous mutation described in the Delta VOC, L452R, was detected at a very low frequency since April 2021 when using a recently described sequence analysis tool (SAM Refiner). Wastewater NGS-based surveillance of

SARS-CoV-2 is a reliable and complementary tool for monitoring the introduction and prevalence of VOCs at a community level allowing early public health decisions. This approach allows the tracking of symptomatic and asymptomatic individuals, who are generally underreported in countries with limited clinical testing capacity. Our results suggests that wastewater-based epidemiology can contribute to improving public health responses in low- and middle-income countries.

Introduction

Early in the COVID-19 pandemic it was found that SARS-CoV-2 was shed in human stool, indicating that human wastewater may be one route by which the virus could be detected. Shortly thereafter, the genetic signal of SARS-CoV-2 was identified in wastewater in many European countries, and in the USA [1, 2]. Since then, SARS-CoV-2 wastewater surveillance (WWS), the measurement of pathogen burden in wastewater by RT-qPCR, has undergone considerable improvement mainly due to an optimization of the concentrating and detecting protocols [3–5]. WWS has been used to evaluate infection trends in the community, monitor public health interventions and to make timely, evidence-based decisions to mitigate the impact of epidemic waves or outbreaks [6]. Additionally, WWS can provide cost-effective support to conventional surveillance methodologies that heavily depend on testing, clinical examination, and case reporting systems.

Few countries in Latin America and the Caribbean regions have developed a WWS program despite the fact that these countries had some of the highest mortality rates. COVID-19 threatens to undo recent gains in human capital outcomes and are the ones that could benefit the most given the low-resource requirements in comparison with other surveillance approaches. WWS can not only provide information on the circulation of pathogens but also identify clusters of cases before levels of sustained transmission are reached, allowing public health professionals to implement targeted control measures [7]. Two years after the first detection of SARS-CoV-2 in wastewater, the World Health Organization (WHO) published guidelines for the WWS of SARS-CoV-2 as a complementary tool to molecular and clinical diagnostics, emphasizing the need of this method to improve surveillance and epidemiological modeling [8]. Since it was first characterized in 2019, SARS-CoV-2 has accumulated mutations that led to the emergence of several viral variants of interest (VOIs) and variants of concern (VOCs). These VOCs accumulated changes in the Receptor-Binding Domain (RBD) of the viral spike (S) protein (D614G, N501Y, E484K, L452R, K417N, and T478K amino acid replacements) that have been associated with an increased transmissibility, a reduction in neutralization by natural or vaccine-derived antibodies and to differences in the clinical outcomes [9–14]. The COVID-19 pandemic has surged in successive epidemic waves, typically associated with the emergence of different VOCs. No short or mid-term public health strategy has been successful in predicting and preventing infection waves [15, 16].

WWS has been used for identifying increasing trends of community-level infections and predicting epidemic peaks with a reasonable anticipation time, thus providing opportune public health interventions [17–21]. Furthermore, by the introduction of genome sequencing methodologies, WWS has expanded from detecting traces of viral RNA by RT-qPCR to tracking the introduction and prevalence of mutations associated with VOCs circulating in the community [22].

Uruguay is a small country in South America with 3 million inhabitants, bordering Brazil (terrestrial borders) and Argentina (maritime borders). The country is organized in 19

departments, with most of the population concentrated in Montevideo, where the capital of the country is located. The intense human mobility across the terrestrial border with Brazil presented a major challenge for public health authorities and clinical data has shown repeated entry of the virus into Uruguay and the subsequent emergence of local outbreaks in Uruguayan border localities [23]. In Uruguay, two major COVID-19 epidemics have been described since the first report of the disease in March 2020. The first wave occurred in April-June 2021 (Fig 1A) and was associated with the Gamma VOC (Fig 1B), while the most recent surge (December-January 2022) was linked to the Omicron VOC. Furthermore, the Delta and other VOIs were also detected [24]. Once the epidemic pattern of SARS-CoV-2 infections was established, characterized by successive waves of distinct VOCs and lineages, a concerted effort was made in Uruguay to establish a uniform program of genomic sequencing of genetic material from nasal swabs from clinical cases, that would allow for the investigation of viral variant distribution and changes over time. Though critically important to establish the epidemiological patterns, a focus on clinical sampling is limited for several reasons, most notably for its cost, which is prohibitive in resource limited communities, and in addition, clinical samples are skewed towards the most symptomatic individuals generating a biased sample [25].

Uruguay established a consistent program of clinical testing that mostly focused on symptomatic patients and their contacts (test-trace-isolate strategy) [26]. Additionally, a program of genomic surveillance investigated the SARS-CoV-2 variant dynamics and distribution over time [24]. The goal of SARS-CoV-2 surveillance using wastewater goes beyond the detection of the genetic signal of the virus it includes monitoring the different viral lineages and their distribution over time. This has been successfully done across the world [27–29] using whole genome sequencing and/or targeted sequencing. Wastewater contains virions shed by numerous infected individuals; therefore, individual mutations cannot be reliably assigned to a specific genome. To overcome these limitations a targeted sequencing approach [28, 30–32] was chosen. Using a similar targeted approach to amplify and sequence a portion of the RBD of SARS-CoV-2 as a single amplicon, we were able to characterize the viral lineages present in sewersheds. Additionally, this strategy allowed us to measure the abundances of the different VOCs present in the community and to identify the introduction of novel variants and VOCs prior to their detection in clinical case-based genomic surveillance.

Materials and methods

A graphical overview of the processing and analysis of the wastewater samples is presented in S1 Fig.

Wastewater sampling and treatment plants

Two different and complementary sampling strategies were carried out. In one, sampling was done once a month from December 2020 to July 2021 (8 samples in total) at a wastewater treatment plant in Rivera, a city with very intense human mobility across the border with Brazil. During the first week of July 2021 wastewater samples were collected in four cities of Uruguay.

The selected cities were: Montevideo, the capital city of Uruguay, with the largest population of the country (about 1.3 million inhabitants); Salto, a border city with Argentina (104011 inhabitants); Rivera, border city with Brazil (64465 inhabitants); and Castillos an eastern seaside city from Rocha department (7541 inhabitants). An additional sample, collected in June 2021 in the Melo city wastewater treatment plant (51830 inhabitants), was also included.

Montevideo city has a combined sewer system that collects rainwater runoff, domestic sewage, and industrial wastewater into one pipe. All the water is transported into a wastewater treatment plant for pre-treatment, then is discharged to an underwater outfall in the Rio de la Plata.
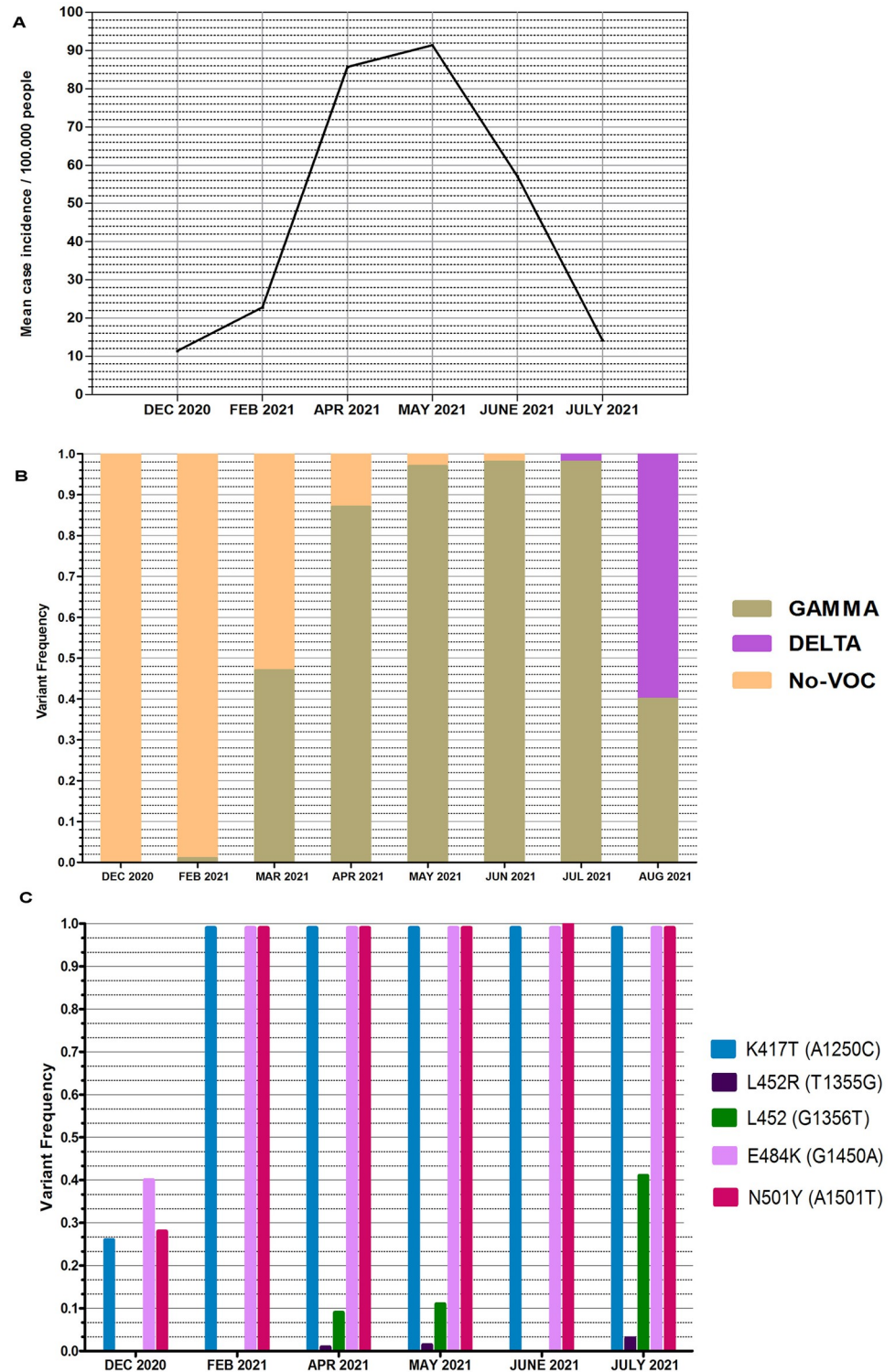
Fig 1. Sampling timeline from Rivera city between December 2020 to July 2021. A) Mean case incidence per 100,000 people from December 2020 to July 2021. **B**) Variant frequency of Gamma, Delta and No-VOC variants in Uruguay between December 2020 and November 2021 according to published epidemiological data. (Ministerio de Salud Pública, 2022). **C**) Variant frequency of the amino acid changes identified in each wastewater sample.

https://doi.org/10.1371/journal.pone.0284483.g001

Salto, Castillos, Melo and Rivera have a different system in which domestic sewage and rainwater runoff do not mix. However, the system is leaky as rainwater does enter the sewage collector and the effluent is consequently diluted when precipitation occurs.

Wastewater concentration

The wastewater concentration method was adapted from a previously published protocol [33]. Briefly, 45 mL of wastewater samples were centrifuged at 4,750 g for 30 minutes. Supernatant was then removed carefully without disturbing the pellet. 40 mL of supernatant was mixed with PEG 8000 (0.1 g/mL) and NaCl (0.0225 g/mL), shaken at room temperature for 15 minutes and centrifuged at 12,000 g (Sigma 2-16KL) for 2 hours at 4˚C. After centrifugation, the supernatant was discarded, and samples were centrifuged again for 5 minutes at 12,000 g to remove the remaining supernatant. Viral pellets were resuspended in 140 µL RNase-free Phosphate Buffered Saline (PBS).

Viral extraction, detection, and quantification

Viral RNA was extracted from wastewater concentrates using the QIAamp® Viral RNA Mini Kit (Qiagen) following manufacturer's instructions. RNA concentrations were determined using a NanoDrop™ Lite Spectrophotometer (Thermo Scientific, USA). SARS-CoV-2 RNAs were detected using the kit COVID-19 RT-qPCR Real TM Ambiental (*Cy5)* (ATGen S.R.L., Montevideo, Uruguay). The kit is a one-step, reverse transcription-Real Time PCR, that targets the N (nucleocapsid) gene of the SARS-CoV-2. Additionally, the kit contains primers and probes for the amplification of an internal control that is added during the RNA extraction step, which constitutes an RNA extraction control. Amplification was performed in a 20 µL reaction mixture containing 10 µL of extracted RNA from each sample. A positive control was provided with the kit, and nuclease free water was used as non-template control. RT-qPCR steps consisted of reverse transcription at 50˚C for 5 min and activation at 95˚C for 20 s, followed by 40 cycles of denaturation at 95˚C for 15 s followed by annealing/extension at 60˚C for 30s. The reactions were performed using a Rotor GeneQ (Qiagen, USA). Real-time PCR Ct value, date and sampling place from Uruguayan wastewater samples are reported in S1 Table.

RBD PCR amplification

The RNAs from the positive samples were reverse transcribed using SuperScript™ IV Reverse Transcriptase and random primers (LifeTechnologies, USA). The RBD region was amplified using Q5® High-Fidelity DNA Polymerase (New England Biolabs, M0491S) [28]. PCR was performed as follows: 98˚C for 1 min, 40 cycles of 98˚C for 15 s, 58˚C for 15 s and 68˚C for 1 min; finally, an extension at 68˚C for 2 min. Amplicons of the expected size (330 bp) were gel-purified with PureLink™ Quick Gel Extraction Kit (Life Technologies, USA).

Ion Torrent next generation sequencing (NGS)

Sequencing libraries were produced from 150 ng purified PCR products using NEBNext® Fast DNA Library Prep Set for Ion Torrent™ (#E6270L, New England Biolabs, Inc) and Ion Xpress™ Barcode Adapters (#4474517, ThermoFisher Scientific). Libraries were quantified using Qubit™ Fluorometer (Qubit™ dsDNA HS Assay Kit), and 50 pM inputs were used for the sequencing template preparation (400bp, Ion Chef protocol) and chip loading in a Ion Chef™ Instrument (#A34019, Ion 510™ & Ion 520™ & Ion 530™ Kit–Chef, ThermoFisher Scientific). Sequencing was performed in the Ion GeneStudio™ S5 System™ at the NGS sequencing

platform of Instituto de Investigaciones Biológicas Clemente Estable (Montevideo, Uruguay). All procedures were according to manufacturer instructions.

NGS data analysis

Analysis of the data obtained included the adapter/quality trimming that was performed with Trim Galore considering base quality >25.

The reference alignment was carried out with the BWA-MEM program and the trimmed reads mapped to two reference sequences of SARS-CoV-2 (Genbank accession number OK091006 and MN908947, corresponding to Delta VOC and Wuhan strain respectively). Consensus calling was obtained with Samtools and IVAR and the spectra of variants in wastewater samples was analyzed with the LoFreq tool [34]. Furthermore, trimmed reads were dereplicated to unique sequences and then mapped to Delta and Wuhan reference sequences of SARS-CoV-2 using BWA-MEM.

To compare the results obtained with an additional approach, mapped RBD amplicons were processed with SAM Refiner [31] to call variants and their abundance.

To identify frequency variants, present in the filtered mapped reads, two requirements were considered: i) sequencing coverage > 100 and ii) variant frequency $\geq$ 5%.

Graphical analysis was performed with GraphPad Prism v5.

iSeq sequencing (NGS)

An aliquot of the cDNA synthesized from RNA isolated from wastewater was further processed for iSeq sequencing of a region within the viral RBD in San Antonio, Texas. This RBD region was amplified using Q5® High-Fidelity DNA Polymerase (New England Biolabs, M0491S) using primers that incorporate Illumina adapters [30]. PCR performed as follows: 98˚C for 30 s plus 40 cycles of 98˚C for 5 s, 53˚C for 15 s and 65˚C for 1 min; finally, an extension of 65˚C for 1 min.

The RBD amplicons were purified using AMPure XP beads (Beckman Coulter, A63881). Index PCR was performed using the Nextera DNA CD Indexes kit (Illumina, 20018707) with 2X KAPA HiFi HotStart ReadyMix (Roche, KK2601), and indexed PCR products purified using AMPure beads (Beckman Coulter, A63881). The indexed libraries were quantified using the Qubit 3.0 and Qubit dsDNA HS Assay Kit (Invitrogen, Q32854) and diluted in 10 mM Tris-HCl to a final concentration of approximately 0.3 ng/µL (1 nM). The libraries were pooled together and diluted to a final concentration of 50 pM. Before sequencing on an Illumina iSeq100, a 10% spike-in of 50 pM PhiX control v3 (Illumina, FC-110-3001) was added to the pooled library. The Illumina iSeq instrument was used to generate paired-end 150 base pair length reads. iSeq reads were uploaded to the BaseSpace Sequence Hub and demultiplexed using a FASTQ generation script. Reads were processed using the published Geneious workflows [35] for preprocessing of NGS reads and assembly of SARS-CoV-2 amplicons variants were called using the Annotate and Predict Find Variations/SNPs in Geneious and verified by using the V-PIPE SARS-CoV-2 application [36]. Paired reads were trimmed, and the adapter sequences removed with the BBDuk plugin. Trimmed reads were aligned to the SARS-CoV-2 reference genome MN908947.

Clinical cases

All clinical data was obtained from public dashboards from the Ministerio de Salud Publica and Instituto Pasteur de Montevideo. All dashboard data is in aggregated format.

Results

Viral RNA detection and RBD Ion Torrent NGS sequencing

After viral concentration, SARS-CoV-2 RNA was detected by RT-qPCR in 12 wastewater samples, with Ct values ranging from 27 to 32.47 (S1 Table). The RBD region amplified in this study included the aa positions 417, 452, 484 and 501 known to be mutated among several VOCs. Substitutions at these positions enabled us to identify SARS-CoV-2 circulating variants in wastewater. A total of 10,849–110,597 sequencing raw reads were obtained for 10 out of 12 samples analyzed. Two samples from Rivera city (from January and March) could not be sequenced. Table 1 summarizes the nucleotide, amino acid changes and their frequencies in the PCR amplicons (compared to Delta reference sequence) using the LoFreq tool. Similar results were obtained by using Wuhan reference sequence NC_045512 for comparison.

The raw sequence data has been deposited in the NCBI-SRA database under the BioSample accession numbers SAMN31036808-SAMN31036817 and BioProject number PRJNA884675.

Analysis of samples from Rivera city

Starting from December 2020, the VOC-associated mutations K417T, E484K and N501Y were detected albeit at a low frequency. However, from February to July 2021 the aa substitutions K417T, E484K and N501Y became the dominant ones with a very high frequency, indicating the presence of the Gamma VOC. Furthermore, in the samples from April and May, a synonymous substitution was identified at position G1356T (L452) with frequencies of 8% and 11%, respectively (Fig 1C). Interestingly, according to the monthly epidemiological reports from the Ministerio de Salud Pública, the Gamma VOC was not detected in clinical specimens until March (Fig 1B).

Analysis of samples for June- July 2021

To take a snapshot of the VOCs distribution across the country for July 2021 samples from four cities (Montevideo, Rivera, Salto and Castillos) were sequenced (Fig 2). An additional sample corresponding to June 2021 from Melo was included.

In all the cases, the three mutations K417T, E484K and N501Y were present at a similar high frequency, above 97%.

Samples from Rivera exhibited greater heterogeneity and had more mutations. Particularly, the synonymous substitution G1356T (L452), which had been detected several months ago, reached a frequency of 41%.

The results of using SAM Refiner to monitor SARS-CoV-2 lineages in wastewater are shown in Table 2. This tool enabled us to further characterize very low frequency variants like the L452R mutation (corresponding to the Delta variant). Samples from Rivera, April, May, and July 2021 contained the L452R amino acid substitution with 1%, 1.4% and 3% abundance respectively.

RBD iSeq NGS sequencing

To further analyze the samples from Uruguay, cDNA was sent to San Antonio, Texas for processing on the iSeq100 pipeline as had been done previously [28]. Challenges in shipping resulted in the package arriving broken and at room temperature. As a result, a wide variety in read quality and quantity was observed for the samples with as few as 54 aligned reads to as many as 47,872 being obtained. Only samples with more than 100 reads were analyzed. It was possible to identify several notable SNPs in the aligning reads including L452R (samples 0 and 7), E484K (samples 0, 7 and 12), N501Y (samples 0, 7 and 12), S494P (sample 0). This

Table 1. Analysis using the LoFreq tool.

Sample Number	Date	Place	Predominant Variant	Nt change (Spike)	AA Change	Variant frequency	Mapped reads	Coverage
13	Dec 2020	Rivera	No VOC*	A1250C	K417T	0.258451	7143	10029
				G1450A	E484K	0.400333		6015
				A1501T	N501Y	0.277101		6402
14	Feb 2021	Rivera	Gamma	A1250C	K417T	0.991176	15137	13712
				G1450A	E484K	0.999021		8175
				A1501T	N501Y	0.999050		8424
15	Apr 2021	Rivera	Gamma	A1250C	K417T	0.986661	5510	5023
				G1356T	None	0.089392		2696
				G1450A	E484K	0.998507		2679
				A1501T	N501Y	0.998625		2910
16	May 2021	Rivera	Gamma	A1250C	K417T	0.989572	4957	4507
				G1356T	None	0.114674		2869
				G1450A	E484K	0.998602		2861
				A1501T	N501Y	0.998326		2987
17	June 2021	Rivera	Gamma	A1250C	K417T	0.990765	4888	4548
				G1450A	E484K	0.999242		2640
				A1501T	N501Y	1.000000		2665
45	June 2021	Melo	Gamma	A1250C	K417T	0.972526	52906	49246
				G1450A	E484K	0.997796		48100
				A1501T	N501Y	0.997755		47654
18	July 2021	Rivera	Gamma	A1250C	K417T	0.994018	6232	5684
				G1356T	None	0.408698		4047
				G1450A	E484K	0.999259		4046
				A1501T	N501Y	0.999282		4176
0	July 2021	Montevideo	Gamma	A1250C	K417T	0.991028	5635	5127
				G1450A	E484K	0.998639		2938
				A1501T	N501Y	0.999684		3160
7	July 2021	Castillos	Gamma	A1250C	K417T	0.977448	3708	3370
				G1356T	None	0.012235		2121
				G1450A	E484K	0.978017		2040
				A1501T	N501Y	0.981375		2094
10	July 2021	Salto	Gamma	A1250C	K417T	0.987458	21690	19455
				G1450A	E484K	0.991361		12038
				A1501T	N501Y	0.992046		12447

The nucleotide and amino acid changes by Ion Torrent sequencing, variant and coverage calculated using the LoFreq tool. The nucleotide and amino acid changes from wastewater amplicons in Uruguayan cities by Ion Torrent sequencing compared to delta reference sequence (Genbank accession number OK091006.1) are shown. No VOC*: AA changes associated with VOC are present at frequencies >5%, however the VOC is not the predominant variant.

approach detected the same amino acid changes that the Ion Torrent NGS sequencing as shown in Table 3.

Discussion

Wastewater detection of SARS-CoV-2 as a component of wastewater-based epidemiology (WWBE) has been demonstrated to be a useful, cost-effective, and reliable tool for monitoring COVID-19 dynamics, the emergence of local viral variants and the introduction of VOCs from other countries [37, 38]. Border crossings presented unique challenges during the
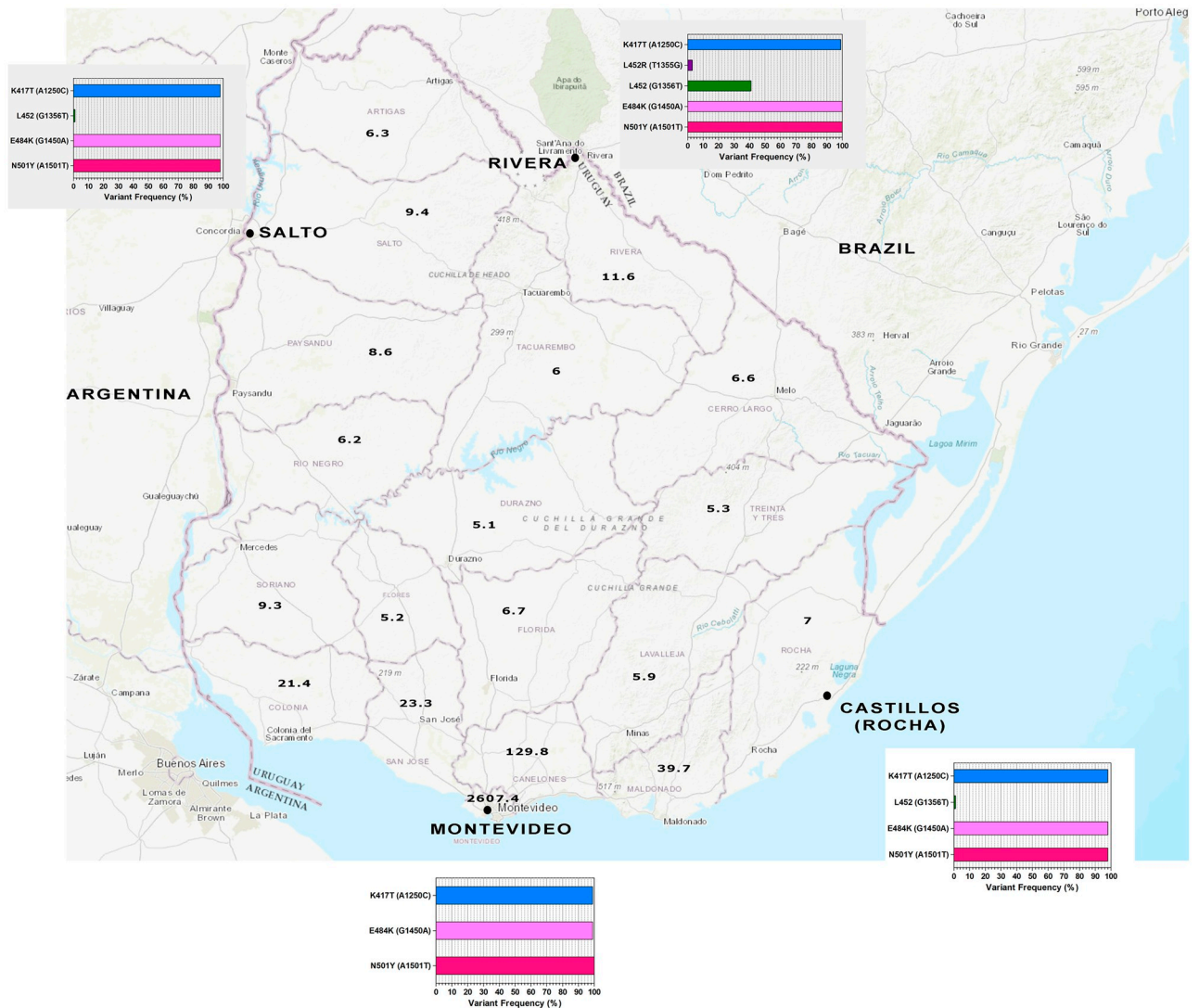
Fig 2. Wastewater sampling from July 2021 in Montevideo, Rivera, Salto, and Castillos. The amino acid changes identified, and their frequencies are shown in graphs for each location. Population density (People/Sq Km) per department are indicated in the map, based on the values extracted from Anuario Estadistico Nacional 2019, 96ª version, Instituto Nacional de Estadística (INE), www.ine.gub.uy. The map of Uruguay was obtained from USGS National Map Viewer (http://viewer.nationalmap.gov/viewer/).

https://doi.org/10.1371/journal.pone.0284483.g002

pandemic and informal border crossings such as those occurring across dry frontiers are hard to capture using clinical testing [39]. Uruguay has a ~1,100 km uninterrupted dry frontier with Brazil and although Uruguayan borders were closed, there was a flow of people across the border [23].

In this study, we monitored the SARS-CoV-2 genetic signal from wastewater in several regions of Uruguay. A targeted-sequencing approach, focused on the RBD of the S protein, allowed us to identify circulating VOCs in different stages of the epidemic outbreak.

In Uruguay, the first COVID-19 case was detected on March 13rd, 2020, and restrictive measures including restriction of social gathering, closing of school and borders were immediately established. Thus, though several local outbreaks occurred, epidemic waves were avoided until October 2021, when cases started to increase steadily [40, 41]. Most of the earlier outbreaks were reported from Rivera, a city bordering Brazil. In June 2021, Brazil was the country

Table 2. Analysis done using the SAMRefiner tool.

Sample Number	Date	Place	Predominant Variant	Nt change (Spike)	AA Change	Variant frequency	Mapped reads	Coverage
13	Dec 2020	Rivera	No VOC*	A1250C	K417T	0.258450	3872	10029
				G1450A	E484K	0.351339		6015
				A1501T	N501Y	0.259874		6402
14	Feb 2021	Rivera	Gamma	A1250C	K417T	0.991175	3421	13712
				G1450A	E484K	0.999021		8175
				A1501T	N501Y	0.999050		8424
15	Apr 2021	Rivera	Gamma	A1250C	K417T	0.986661	1465	5023
				T1355G	L452R	0.009637		2698
				G1356T	None	0.089392		2696
				G1450A	E484K	0.998507		2679
				A1501T	N501Y	0.998625		2910
16	May 2021	Rivera	Gamma	A1250C	K417T	0.989572	1560	4507
				T1355G	L452R	0.013598		2868
				G1356T	None	0.114674		2869
				G1450A	E484K	0.998602		2861
				A1501T	N501Y	0.998326		2987
17	June 2021	Rivera	Gamma	A1250C	K417T	0.990765	1384	4548
				G1450A	E484K	0.999242		2640
				A1501T	N501Y	1.000000		2665
45	June 2021	Melo	Gamma	A1250C	K417T	0.972526	14513	49246
				G1450A	E484K	0.998025		48100
				A1501T	N501Y	0.997755		47654
18	July 2021	Rivera	Gamma	A1250C	K417T	0.994018	2059	5684
				T1355G	L452R	0.032362		4048
				G1356T	None	0.408698		4047
				G1450A	E484K	0.999258		4045
				A1501T	N501Y	0.999282		4176
0	July 2021	Montevideo	Gamma	A1250C	K417T	0.991028	1607	5127
				G1450A	E484K	0.998979		2938
				A1501T	N501Y	0.999683		3160
7	July 2021	Castillos	Gamma	A1250C	K417T	0.977448	1332	3370
				G1356T	None	0.014131		2123
				G1450A	E484K	0.963890		2047
				A1501T	N501Y	0.981375		2094
10	July 2021	Salto	Gamma	A1250C	K417T	0.987458	5249	19455
				G1450A	E484K	0.991360		12038
				A1501T	N501Y	0.992046		12447

Nucleotide and amino acid changes by Ion Torrent sequencing, variant and coverage calculated using the SAMRefiner tool. The nucleotide and amino acid changes from wastewater amplicons in Uruguayan cities by Ion Torrent sequencing compared to delta reference sequence (Genbank accession number OK091006.1) are shown. No VOC*: AA changes associated with VOC are present at frequencies >5%, however the VOC is not the predominant variant.

with the second-highest rate of deaths per million and the fourth by the number of cases per million in Latin America and the Caribbean [42, 43].

Immediately, this city became the main concern for the government, and it was identified as a potential route for the introduction of active, imported COVID-19 cases and SARS-CoV-2 lineages.

Table 3. Analysis using Geneious.

Sample Number	Date	Place	Predominant Variant	Nt change (Spike)	AA Change	Variant frequency	Mapped reads	Coverage
0	July 2021	Montevideo	No VOC*	T1355G	L452R	0.117	47872	24310
				G1450A	E484K	0.062		23064
				T1480C	S494P	0.115		23227
				A1501T	N501Y	0.28		23014
7	July 2021	Castillos	Gamma	T1355G	L452R	0.031	2466	1210
				G1450A	E484K	0.956		1229
				A1501T	N501Y	0.96		1224
12	July 2021	Cerro Largo	No VOC*	G1450A	E484K	0.174	38121	18429
				A1501T	N501Y	0.435		18382

Nucleotide and amino acid changes by Illumina iSeq sequencing, variant and coverage calculated using the Annotate and Predict Find Variations/SNPs in Geneious and verified by using the V-PIPE SARS-CoV-2 application (36). The nucleotide and amino acid changes from wastewater in Uruguayan cities by Illumina iSeq sequencing compared to Wuhan reference sequence (Genbank accession number MN908947) are shown.

No VOC*: AA changes associated with VOC are present at frequencies >5%, however the VOC is not the predominant variant.

https://doi.org/10.1371/journal.pone.0284483.t003

In Rivera, qPCR data showed that Ct values in wastewater samples remained quite stable from December to July 2021, despite the cases/million people incidence rate exhibiting clear fluctuations in Uruguay during that period. This discrepancy can be explained by the lack of a reliable internal indicator to calculate the viral RNA recovery efficiency from wastewater. Limitations of this study also include the population mobility that characterizes border towns that are always difficult to measure even more during the implementation of pandemic restrictions.

The results from the targeted amplicon-NGS sequencing of the December samples from Rivera showed early evidence of the presence, though at low frequency, of several mutations associated to the Gamma VOC (K417T, E484K and N501Y), which were then confirmed in February when this variant became dominant in wastewater in that city. The Gamma VOC was identified in Manaos, Brazil, in November 2020 and then spread throughout the country within weeks [44]. Our findings agree with recent reports, based on phylogeographic analyses, that propose the outbreaks in Uruguay were mostly caused by introductions from neighboring South American countries, particularly Brazil [45, 46]. During this time an additional variant, named P.2 and classified as a VOI, was introduced to Uruguay, and detected in clinical swabs. Of note, this variant shared the E484K mutation with the Gamma variant [47], so they would not be differentiated using our approach.

Interestingly, the epidemiological data from the nasopharyngeal swabs samples in Uruguay revealed that the Gamma VOC was not detected in human cases until March. However, it was detected in wastewater in Rivera city, albeit at a low frequency; and by April it had become dominant (Fig 1C). Furthermore, a Bayesian analysis reported by Rego et al. [46] estimated the median time of the most recent common ancestor of the Gamma clades identified in COVID-19 cases in Uruguay to be mid-February to early March 2021.

Even though the Gamma VOC fueled the first epidemic wave in Uruguay, additional viral mutations were detected in human clinical samples that could not be detectable/identified by our targeted amplicon-approach. However, the exact role of these variants in the epidemic dynamics is unclear since they were rapidly displaced by Gamma [48].

In late 2020, the B.1.617 VOC (named Delta) was first detected in India and by July 2021 it had been reported in 60 countries [49]. To assess whether this VOC was already circulating in our country, wastewater was sampled in the first week of July from 4 regions. No evidence of

the Delta VOC associated mutations were detected. However, a more detailed analysis performed with SAM Refiner, a recently developed bioinformatic tool, showed that the L452R mutation was indeed present by April 2021 in Rivera (Table 2). Although we set a cut-off criterion of variant frequency $\geq$ 5%, we decided to show these data as it is considered a Delta VOC marker [50]. Interestingly, according to reports from the Ministerio de Salud Pública, the first clinical case associated with Delta infection dated from mid-July, and by August circulation at the community level had not been observed yet [51].

In this work, Spike targeted-sequencing enabled the early identification of SARS-CoV-2 VOCs circulation in wastewater samples, even at low frequencies and prior in advance of their detection in clinical samples.

Even though the number of mapped reads were modest, the results of VOCs variant frequencies and counts obtained by both bioinformatic approaches highly correlated, indicating a robust and reliable analysis.

Frequency analyses evidenced the rapid substitution of No-VOC lineages by the Gamma variant during the December-February 2020 period. The Gamma VOC likely originated in Rivera city and subsequently spread to the rest of the territory. Our analysis did not identify any variants/mutations in wastewater that had not been reported in clinical cases.

In summary, this study supports the use of wastewater surveillance as a reliable and complementary tool for monitoring the introduction, prevalence and spread of VOC and VOI at the community level. In particular, the early detection of new variants, before their detection in patients, can help local authorities to make informed public health decisions supporting better health outcomes.

Supporting information

S1 Fig. Graphical overview of the processing and analysis of the wastewater samples.
(TIF)

S1 Table. Real-time PCR Ct value, date, and sampling place from Uruguayan wastewater samples.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Davida S. Smyth, Santiago Mirazo, Mónica Trujillo.

Data curation: Florencia Cancela, Natalia Ramos, Davida S. Smyth, Germán Botto, Starla G. Thornhill, Santiago Mirazo.

Formal analysis: Florencia Cancela, Natalia Ramos, Caterina Rufo, Starla G. Thornhill, Mónica Trujillo.

Funding acquisition: Davida S. Smyth, Mónica Trujillo.

Investigation: Florencia Cancela, Davida S. Smyth, Claudia Etchebehere, Mabel Berois, Jesica Rodríguez, Caterina Rufo, Alicia Alemán, Liliana Borzacconi, Julieta López, Elizabeth González, Germán Botto, Santiago Mirazo, Mónica Trujillo.

Methodology: Davida S. Smyth, Santiago Mirazo, Mónica Trujillo.

Supervision: Davida S. Smyth, Santiago Mirazo.

Writing – original draft: Davida S. Smyth, Santiago Mirazo, Mónica Trujillo.

Writing – review & editing: Florencia Cancela, Davida S. Smyth, Mabel Berois, Starla G. Thornhill, Santiago Mirazo, Mónica Trujillo.

References

1. Kirby AE, Walters MS, Jennings WC, et al. Using Wastewater Surveillance Data to Support the COVID-19 Response—United States, 2020–2021. *MMWR Morb Mortal Wkly Rep*. 2021; 70(36):1242–1244. Published 2021 Sep 10. https://doi.org/10.15585/mmwr.mm7036a2 PMID: 34499630

2. Wu F, Lee WL, Chen H, et al. Making waves: Wastewater surveillance of SARS-CoV-2 in an endemic future. *Water Res*. 2022; 219:118535. https://doi.org/10.1016/j.watres.2022.118535 PMID: 35605390

3. Ahmed W, Simpson SL, Bertsch PM, Bibby K, Bivins A, Blackall LL, et al. Minimizing errors in RT-PCR detection and quantification of SARS-CoV-2 RNA for wastewater surveillance. *Sci. Total Environ*. 2022; 805. https://doi.org/10.1016/j.scitotenv.2021.149877 PMID: 34818780

4. Pérez-Cataluña A, Cuevas-Ferrando E, Randazzo W, Falcó I, Allende A, Sánchez G. Comparing analytical methods to detect SARS-CoV-2 in wastewater. *Sci. Total Environ*. 2021; 758; 143870. https://doi.org/10.1016/j.scitotenv.2020.143870 PMID: 33338788

5. Crits-Christoph A, Kantor RS, Olm MR, et al. Genome Sequencing of Sewage Detects Regionally Prevalent SARS-CoV-2 Variants. *mBio*. 2021; 12(1):e02703–20. https://doi.org/10.1128/mBio.02703-20 PMID: 33468686

6. Kumar M, Jiang G, Kumar Thakur A, Chatterjee S, Bhattacharya T, Mohapatra S, et al. Lead time of early warning by wastewater surveillance for COVID-19: Geographical variations and impacting factors. *Chem Eng J*. 2022 Aug 1; 441:135936. https://doi.org/10.1016/j.cej.2022.135936 PMID: 35345777

7. The World Bank. Wastewater testing can improve and enhance the response to COVID-19 in Latin America and the Caribbean. 2022, Jan 25. (https://www.worldbank.org/en/news/press-release/2022/01/24/pruebas-de-aguas-residuales-pueden-mejorar-respuesta-contra-covid19-america-latina-y-caribe).

8. World Health Organization. Environmental surveillance for SARS-COV-2 to complement public health surveillance–Interim Guidance. 2022. (https://www.who.int/publications-detail-redirect/WHO-HEP-ECH-WSH-2022.1).

9. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature*. 2021; 592: 116–121. https://doi.org/10.1038/s41586-020-2895-3 PMID: 33106671

10. Tian F, Tong B, Sun L, Shi S, Zheng B, Wang Z, et al. N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2. eLife. 2021; Aug 20; 10: e69091. https://doi.org/10.7554/eLife.69091 PMID: 34414884

11. Jangra S, Ye C, Rathnasinghe R, Stadlbauer D, Personalized Virology Initiative study group, Krammer F, et al. SARS-CoV-2 spike E484K mutation reduces antibody neutralization. *Lancet Microbe*. 2021; Jul; 2(7):e283–e284.

12. Tchesnokova V, Kulasekara H, Larson L, Bowers V, Rechkina E, Kisiela D, et al. Acquisition of the L452R Mutation in the ACE2-Binding Interface of Spike Protein Triggers Recent Massive Expansion of SARS-CoV-2 Variants. *J Clin Microbiol*. 2021 Oct 19; 59(11):e0092121. https://doi.org/10.1128/JCM.00921-21 PMID: 34379531

13. Fratev F. N501Y and K417N Mutations in the Spike Protein of SARS-CoV-2 Alter the Interactions with Both hACE2 and Human-Derived Antibody: A Free Energy of Perturbation Retrospective Study. *J Chem Inf Model*. 2021 Dec 27; 61(12):6079–6084. https://doi.org/10.1021/acs.jcim.1c01242 PMID: 34806876

14. Di Giacomo S, Mercatelli D, Rakhimov A, Giorgi FM. Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Spike mutation T478K. *J Med Virol*. 2021 Sep; 93(9):5638–5643. https://doi.org/10.1002/jmv.27062 PMID: 33951211

15. Tatsi EB, Filippatos F, Michos A. SARS-CoV-2 variants and effectiveness of vaccines: a review of current evidence. *Epidemiol Infect*. 2021 Nov 4; 149:e237. https://doi.org/10.1017/S0950268821002430 PMID: 34732275

16. Jung C, Kmiec D, Koepke L, Zech F, Jacob T, Sparrer KMJ, et al. Omicron: What Makes the Latest SARS-CoV-2 Variant of Concern So Concerning? *J Virol*. 2022 Mar 23; 96(6):e0207721. https://doi.org/10.1128/jvi.02077-21 PMID: 35225672

17. Center for Disease Control and Prevention -CDC. National Wastewater Surveillance System (NWSS). 2022. (https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html

18. Sharara N, Endo N, Duvallet C, Ghaeli N, Matus M, et al. Wastewater network infrastructure in public health: Applications and learnings from the COVID-19 pandemic. *PLOS Global Public Health.* 2021; 1 (12): e0000061. https://doi.org/10.1371/journal.pgph.0000061 PMID: 34927170

19. Ahmed W, Tscharke B, Bertsch PM, Bibby K, Bivins A, Choi P, et al. SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission in the community: A temporal case study. *Sci Total Environ.* 2021; Mar 20; 761:144216. https://doi.org/10.1016/j.scitotenv.2020.144216 PMID: 33360129

20. Proverbio D, Kemp F, Magni S, Ogorzaly L, Cauchie HM, Gonçalves J, et al. Model-based assessment of COVID-19 epidemic dynamics by wastewater analysis. *Sci Total Environ.* 2022; Jun 25; 827:154235. https://doi.org/10.1016/j.scitotenv.2022.154235 PMID: 35245552

21. Joseph-Duran B, Serra-Compte A, Sàrrias M, Gonzalez S, López D, Prats C, et al. Assessing wastewater-based epidemiology for the prediction of SARS-CoV-2 incidence in Catalonia. *Sci Rep.* 2022; Sep 5; 12(1):15073. https://doi.org/10.1038/s41598-022-18518-9 PMID: 36064874

22. Cutrupi F, Cadonna M, Manara S, Foladori P. Surveillance of SARS-CoV-2 in extensive monitoring of municipal wastewater: key issues to yield reliable results. *Water Sci Technol.* 2021; Dec; 84(12):3508–3514. https://doi.org/10.2166/wst.2021.469 PMID: 34928822

23. Mir D, Rego N, Resende PC, Tort F, Fernández-Calero T, Noya V, et al. Recurrent Dissemination of SARS-CoV-2 Through the Uruguayan-Brazilian Border. *Front Microbiol.* 2021; May 28; 12:653986. https://doi.org/10.3389/fmicb.2021.653986 PMID: 34122369

24. Working Group on Genomic Surveillance of SARS-CoV-2. Institut Pasteur de Montevideo. 2022. https://pasteur.uy/en/news/one-year-monitoring-sars-cov-2-variants/.

25. Shrestha S, Yoshinaga E, Chapagain SK, Mohan G, Gasparatos A, Fukushi K. Wastewater-Based Epidemiology for Cost-Effective Mass Surveillance of COVID-19 in Low- and Middle-Income Countries: Challenges and Opportunities. *Water.* 2021; 13(20):2897. https://doi.org/10.3390/w13202897.

26. Hernández D, Lòpez C. COVID-19 en Uruguay: estrategias de abordaje y dinámicas políticas. Serie Documentos RISEP, 2021. No24.

27. Karthikeyan S, Levy JI, De Hoff P, Humphrey G, Birmingham A, Jepsen K, et al. Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature.* 2022 Sep; 609(7925):101–108. https://doi.org/10.1038/s41586-022-05049-6 PMID: 35798029

28. Smyth DS, Trujillo M, Gregory DA, Cheung K, Gao A, Graham M, et al. Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nat Commun.* 2022; Feb 3; 13(1):635. https://doi.org/10.1038/s41467-022-28246-3 Erratum in: *Nat Commun.* 2022; Mar 30;13(1):1836. PMID: 35115523

29. Amman F, Markt R, Endler L, Hupfauf S, Agerer B, Schedl A. Viral variant-resolved wastewater surveillance of SARS-CoV-2 at national scale. *Nat Biotechnol.* 2022 Jul 18. https://doi.org/10.1038/s41587-022-01387-y PMID: 35851376

30. Bewicke-Copley F, Arjun Kumar E, Palladino G, Korfi K, Wang J. Applications and analysis of targeted genomic sequencing in cancer studies. *Comput Struct Biotechnol J.* 2019; Nov 7; 17:1348–1359. https://doi.org/10.1016/j.csbj.2019.10.004 PMID: 31762958

31. Gregory DA, Wieberg CG, Wenzel J, Lin CH, Johnson MC. Monitoring SARS-CoV-2 Populations in Wastewater by Amplicon Sequencing and Using the Novel Program SAM Refiner. *Viruses.* 2021; Aug 19; 13(8):1647. https://doi.org/10.3390/v13081647 PMID: 34452511

32. Gregory DA, Trujillo M, Rushford C, Flury A, Kannoly S, San KM, et al. Genetic diversity and evolutionary convergence of cryptic SARS- CoV-2 lineages detected via wastewater sequencing. *PLoS Pathog.* 2022; Oct 14; 18(10):e1010636. https://doi.org/10.1371/journal.ppat.1010636 PMID: 36240259

33. Wu F, Zhang J, Xiao A, Gu X, Lee WL, Armas F, et al. SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases. *mSystems.* 2020 Jul 21; 5(4): e00614–20. https://doi.org/10.1128/mSystems.00614-20 PMID: 32694130

34. Wilm A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, Wong CH, et al. LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res.* 2012; 40: 11189–11201. https://doi.org/10.1093/nar/gks918 PMID: 23066108

35. Miller H. Geneious Knowledge Base. 2021. https://help.geneious.com/hc/en-us/articles/360045070991-Assembly-of-SARS-CoV-2-genomes-from-tiled-amplicon-Illumina-sequencing-using-Geneious-Prime).

36. Posada-Céspedes S, Seifert D, Topolsky I, Jablonski KP, Metzner KJ, Beerenwinkel N. V-pipe: a computational pipeline for assessing viral genetic diversity from high-throughput data. *Bioinformatics*. 2021; Jan 20; 37(12):1673–1680. https://doi.org/10.1093/bioinformatics/btab015 PMID: 33471068

37. Lodder W, de Roda Husman AM. SARS-CoV-2 in wastewater: potential health risk, but also data source. Lancet Gastroenterol *Hepatol*. 2020; Jun; 5(6):533–534. https://doi.org/10.1016/S2468-1253(20)30087-X PMID: 32246939

38. Randazzo W, Cuevas-Ferrando E, Sanjuán R, Domingo-Calap P, Sánchez G. Metropolitan wastewater analysis for COVID-19 epidemiological surveillance. *Int J Hyg Environ Health*. 2020; Sep; 230:113621. https://doi.org/10.1016/j.ijheh.2020.113621 PMID: 32911123

39. Miller MA, Astuti R, Hirsch P, Marschke M, Rigg J, Saksena-Taylor P, et al. Selective border permeability: Governing complex environmental issues through and beyond COVID-19. *Polit Geogr*. 2022; Aug; 97:102646. https://doi.org/10.1016/j.polgeo.2022.102646 PMID: 35342230

40. Ministerio de Salud Pública. Informe epidemiológico COVID-19 al 30 de octubre de 2020. (https://www.gub.uy/ministerio-salud-publica/comunicacion/noticias/informe-epidemiologico-covid-19-del-6-noviembre-2020)

41. Panzera Y, Goñi N, Calleros L, Ramos N, Frabasile S, Marandino A, et al. Genome Sequences of SARS-CoV-2 P.1 (Variant of Concern) and P.2 (Variant of Interest) Identified in Uruguay. *Microbiol Resour Announc*. 2021; May 27; 10(21):e0041021. https://doi.org/10.1128/MRA.00410-21 PMID: 34042476

42. World Bank. 2022. Brazil–COVID-19 in Latin America and Caribbean: High Frequency Phone Surveys —Results Phase Two, Wave One. 2021. Washington, DC. (https://openknowledge.worldbank.org/handle/10986/37541).

43. Diaz RS, Vergara TRC. The COVID-19 second wave: A perspective to be explored. *Braz J Infect Dis*. 2021; Jan-Feb; 25(1):101537. https://doi.org/10.1016/j.bjid.2020.101537 PMID: 33378685

44. Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, et al. The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat Rev Genet*. 2021 Dec; 22(12):757–773. https://doi.org/10.1038/s41576-021-00408-x PMID: 34535792

45. Elizondo V, Harkins GW, Mabvakure B, Smidt S, Zappile P, Marier C, et al. SARS-CoV-2 genomic characterization and clinical manifestation of the COVID-19 outbreak in Uruguay. *Emerg Microbes Infect*. 2021; Dec; 10(1):51–65. https://doi.org/10.1080/22221751.2020.1863747 PMID: 33306459

46. Rego N, Costábile A, Paz M, Salazar C, Perbolianachis P, Spangenberg L, et al. Real-Time Genomic Surveillance for SARS-CoV-2 Variants of Concern, Uruguay. *Emerg Infect Dis*. 2021a; Nov; 27 (11):2957–2960. https://doi.org/10.3201/eid2711.211198 PMID: 34437831

47. Lamarca AP, de Almeida LGP, Francisco RDS Jr, Lima LFA, Scortecci KC, Perez VP, et al. Genomic surveillance of SARS-CoV-2 tracks early interstate transmission of P.1 lineage and diversification within P.2 clade in Brazil. *PLoS Negl Trop Dis*. 2021; Oct 13; 15(10):e0009835. https://doi.org/10.1371/journal.pntd.0009835 PMID: 34644287

48. Rego N, Salazar C, Paz M, Costábile A, Fajardo A, Ferrés I, et al. Emergence and Spread of a B.1.1.28-Derived P.6 Lineage with Q675H and Q677H Spike Mutations in Uruguay. *Viruses*. 2021b; Sep 10; 13(9):1801. https://doi.org/10.3390/v13091801 PMID: 34578382

49. CDC. COVID-19: SARS-CoV-2 variant classifications and definitions. Atlanta, GA: US Department of Health and Human Services, CDC; 2021.

50. Hajizadeh F, Khanizadeh S, Khodadadi H, Mokhayeri Y, Ajorloo M, Malekshahi A, et al. SARS-COV-2 RBD (Receptor binding domain) mutations and variants (A sectional-analytical study). *Microb Pathog*. 2022 Jul; 168:105595. https://doi.org/10.1016/j.micpath.2022.105595 PMID: 35597364

51. Ministerio de Salud Pública. Casos activos de nuevas variantes de SARS-CoV-2 en Uruguay al día de hoy. 2021 Aug 6. (https://www.gub.uy/ministerio-salud-publica/sites/ministerio-saludpublica/files/documentos/noticias/Variantes%20activas%20hoy%20-%2006-08-2021.pdf).