NATIONAL INSTITUTE FOR COMMUNICABLE DISEASES
Division of the National Health Laboratory Service

WASTEWATER-BASED EPIDEMIOLOGY FOR SARS-CoV-2 SURVEILLANCE IN SOUTH AFRICA

Detection, quantitation and genomic sequencing at sentinel sites in South Africa, March 2021- April 2023
WEEK 16 2023

Co-funded by the Water Research Commission, the Bill and Melinda Gates Foundation and the NICD

Chinwe Iwu-Jaja[1*], Setshaba Taukobong[1*], Said Rachida[1], Nkosenhle Ndlovu[1], Mokgaetji Macheke[1], Wayne Howard[1], Shelina Moonsamy[1], Gina Pocock[3], Leanne Coetzee[3], Janet Mans[4], Lisa Schaefer[5], Wouter J. Le Roux[5], Annancietar Gomba[6], Don Jambo[6], David Moriah de Villiers[7], Nadine Lee Lepart[7], Shaun Groenink[8], Neil Madgwick[9], Martie van der Walt[10], Awelani Mutshembele[10], Leanne Pillay[11], Faizal Bux[11], Isaac Dennis Amoah[11], Natacha Berkowitz[12], Jay Bhagwan[12], Melinda Suchard[1,14], Kerrigan McCarthy[#1,15], Mukhlid Yousif[#1,16] for the South African Collaborative COVID-19 Environmental Surveillance System (SACCESS) network.

[1]Centre for Vaccines and Immunology, National Institute for Communicable Diseases, a division of the National Health Laboratory Service, South Africa
[3]Waterlab, (Pty) Ltd, Pretoria
[4]Department of Medical Virology, University of Pretoria
[5]Water Centre, Council for Scientific and Industrial Research (CSIR), Pretoria
[6]National Institute for Occupational Health, a division of the National Health Laboratory Service, Johannesburg
[7]Lumegen Laboratories, (Pty) Ltd, Potchefstroom
[8] Greenhill Laboratories
[9]Praecautio
[10]Tuberculosis Platform, South African Medical Research Council, Pretoria.
[11]Institute of Wastewater Management, Durban University of Technology
[12]City of Cape Town Health Department
[13] Water Research Commission, Pretoria
[14]Department of Chemical Pathology, School of Pathology, University of the Witwatersrand, Johannesburg
[15]School of Public Health, University of the Witwatersrand, Johannesburg
[16]Department of Virology, School of Pathology, University of the Witwatersrand, Johannesburg

*joint first authors
[#]joint last authors

OVERVIEW

This report summarises and interprets findings from detection, quantification and sequencing of SARS-CoV-2 by the National Institute for Communicable Diseases (NICD) Centre for Vaccines and Immunology from influent (untreated) wastewater in 17 wastewater treatment plants (WWTPs) across five South African provinces. Levels of SARS-CoV-2 in wastewater correlate with population levels of SARS-CoV-2 over time and indicate the geographic distribution of disease. Variants of SARS-CoV-2 can be identified in wastewater through detection of single-nucleotide polymorphisms (SNPs) that are specific to each variant. These variants are shown to correspond to variants prevalent in clinical cases, across time and place. SARS-CoV-2 is shed from symptomatic and asymptomatic persons in stool but is not transmitted by faecal-oral route nor via wastewater. This report is based on data collected from June 2021 until 19 April 2023 (Epidemiological week 16, 2023). Results from wastewater testing should be read and interpreted together with testing and genomic reports generated by the Centre for Respiratory Diseases and Meningitis found at https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/

- Part 1 of this report presents methods and results of quantitative testing of wastewater.
- Part 2 of this report presents methods and results from sequencing of SARS-CoV-2 RNA fragments in wastewater.

SARS-CoV-2 levels in wastewater are mostly low to moderate across the country. Sequencing data from week 13, 2023 show that recombinant lineages XBB.1.9, XBB.1.9.1 with XBB.1.5 dominance are circulating in March, 2023 in Daspoort, Rooiwal, Central eThekwini, Northern eThekwini, Hartbeesfontein and Bloemspruit. The predominant lineages circulating in clinical samples in the recent week are XBB.1.5 and XBB sub-lineages. The significance of this emergence is not yet known, however, lineage XBB.1.9.1 is currently increasing and circulating in Indonesia, South East Asia and Europe. Detailed analyses are described below.

HIGHLIGHTS – sample collection dates up to 19 April, 2023 (Epi week 16)

SARS-CoV-2 levels in wastewater:

- The levels of SARS-CoV-2 in wastewater treatment plants across the country were mostly moderate and others low.
- Moderate levels were detected in WWTPs tested from Tshwane district, Ekurhuleni metro, eThekwini and Nelson Mandela district.
- Low levels of SARS-CoV-2 were detected in WWTPs tested from Mangaung metro and City of Cape Town.
- These results are based solely on sites tested by the NICD.

*Note: The presence and increase/decrease of SARS-CoV-2 RNA in wastewater signify ongoing and increasing/decreasing transmission of the virus amongst populations that are serviced by particular sewer networks. The determination of a resurgence (or 'wave') of SARS-CoV-2 is made through evaluation of clinical testing data (including numbers of positive tests, percentage testing positive), hospitalisation and mortality data.

SARS-CoV-2 genomics in wastewater:

Sequencing data available up to week 09 (**28th March, 2023**) shows that recombinant lineages XBB.1.9, XBB.1.9.1 with XBB.1.5 dominance are overall, circulating in March, 2023 in South Africa.

PART 1: Detection and quantification of SARS-CoV-2 in wastewater

Background

The detection and monitoring of SARS-CoV-2 through wastewater was first proposed in April 2020. Initial reports describing the feasibility and practical usefulness of this approach emerged simultaneously from several countries during August 2020. Recent evidence has shown that SARS-CoV-2 can be detected in wastewater prior to the appearance of clinical cases, and longitudinal tracking of SARS-CoV-2 viral load in wastewater correlates with the burden of clinically diagnosed cases. Furthermore, the sequencing of SARS-CoV-2 RNA fragments in wastewater has identified variants of concern as well as mutations not detected in clinical cases.

In South Africa, SARS-CoV-2 epidemiology is monitored through laboratory testing of clinical cases using reverse-transcriptase polymerase chain reaction (RT-PCR) tests and rapid antigen tests, COVID-19 hospital admissions and COVID-19 related deaths. Laboratory testing data is sent by testing laboratories to the National Institute for Communicable Diseases (NICD) via the DATCOV system. From these data sources, epidemiological indicators including incidence rates of testing and case detection, hospitalisation and death rates are made available to key stakeholders and the general public.

Clinical epidemiology based on reporting of laboratory-confirmed cases of SARS-CoV-2 has limitations. Household transmission studies in South African urban and rural settings have demonstrated that a large proportion of cases are asymptomatic or so mild as not to elicit health-seeking, and that laboratory-confirmed cases likely represent less than 10% of SARS-CoV-2 cases prevalent in a community at any given time. Secondly, there is increasing use of rapid antigen detection tests in clinical settings. Results of these tests may not be reported to surveillance networks. Consequently, laboratory diagnosis is increasingly less representative of the burden of disease.

Methods

Outbreak context and clinical case epidemiology

Since the first case of SARS-CoV-2 in South Africa was detected on 3rd March 2020, laboratories in the country have conducted **over 25 million RT-PCR and antigen tests**. Five distinct waves of SARS-CoV-2 infection have occurred so far, peaking in June 2020, December 2020, July 2021, December 2021 and June 2022, respectively. The current de-duplicated and geospatially allocated national line list of laboratory-confirmed cases of SARS-CoV-2 (identified by RT-PCR or antigen test) is provided by the NICD for comparison with results from SARS-CoV-2 testing of wastewater.

Establishment of the laboratory testing network

Commencing in 2018, the NICD had been conducting testing of wastewater for poliovirus as part of the National Department of Health's polio surveillance programme. In 2020, the NICD commenced testing of influent wastewater samples from these 18 sites, including eight in Gauteng Province, two in the City

of Cape Town (Western Cape Province), two in Mangaung (Free State Province), two in eThekwini (KwaZulu- Natal Province) and four in Eastern Cape Province (two in Buffalo City Metro and two in Nelson Mandela Metro). Quantitative testing results for these sites are available from week 8 of 2021, onwards.

SARS-CoV-2 detection and quantitation methodology

The general approach of SARS-CoV-2 detection in wastewater used at all network laboratories is virus concentration, followed by nucleic acid extraction and molecular detection. At the identified wastewater treatment facilities grab or passive samples of influent are collected and transported at <5°C to the testing facility. Table 1 summarises the sample collection, processing and detection methodology used by laboratories involved in the surveillance project. The levels of SARS-CoV-2 in wastewater are reported in copies/mL of wastewater.

Table 1. Sampling and methodology used by laboratories involved in the NICD-WRC led COVID-19 wastewater surveillance project.

Name of laboratory	Sampling	Virus concentration	Nucleic acid extraction	Molecular analysis	Molecular analysis platform
National Institute for Communicable Diseases (NICD)	Grab	Ultrafiltration (Centricon® Plus-70 centrifugal ultra-filter device)	QIAamp® viral RNA mini kit	RT-qPCR[a] using the Allplex™ 2019-nCoV Assay and the EDX SARS-CoV-2 standard	7500 Real-Time PCR System (Applied Biosystems

Interpretation of SARS-CoV-2 levels in wastewater

Interpretation of SARS-CoV-2 wastewater levels is evolving. We have elected to use interpretive principles outlined in Table 2 to support public health preparedness and response activities. In general, increasing or decreasing trends in levels are reported based on two or more results, as a single sample that increases or decreases compared with the result from the previous week may represent an outlier. Small changes (up to 0.5 log copies/ml) are not regarded as significant changes unless they form part of a general upward or downward trend. Comparison of results over time when quantification is done by the same laboratory using the same quantitative methodology is meaningful. The use of different methodologies by different laboratories precludes comparison of quantitative results across laboratories. The Ct values is an alternative for quantification. Changes in the Ct value of SARS-CoV-2 give an indication of whether the burden of disease is increasing or decreasing.

Table 2. Principles of SARS-CoV-2 detection and quantification on influent samples from wastewater treatment plants and interpretive principles to guide application of test results to support COVID-19 public health responses, South Africa.

Testing modality	Interpretive principles to support public health responses
Detection of SARS-CoV-2	When a test result changes from • positive to negative, this signifies fewer/no cases in population • negative to positive, this indicates the need for increased population awareness and action • Qualitative results (presence or absence) are comparable between laboratories • Changes in the Cycle threshold (Ct) value of SARS-CoV-2 give an indication of whether the burden of disease is increasing or decreasing
Quantification of SARS-CoV-2	• The concentration of SARS-CoV-2 at a particular facility may be used to infer the burden of SARS-CoV-2 in the population served by the wastewater treatment facility. • Changes in the concentration of SARS-CoV-2 give an indication of whether the burden of disease is increasing or decreasing • Quantitative results between laboratories are not comparable. Quantitative results should be interpreted for a single wastewater treatment plant tested by the same laboratory using the same methodology over time.

Results

Gauteng Province

A: City of Tshwane

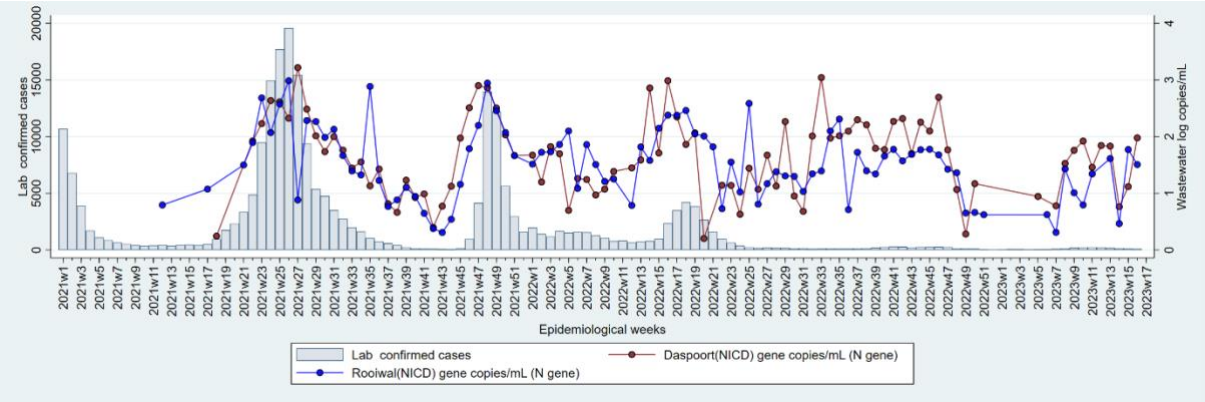


Figure 1A. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) for selected wastewater treatment plants (WWTP) and metropolitan areas in Tshwane District Municipality (Tshwane North), Gauteng Province during epidemiological weeks 1 of 2021 to week 16 of 2023.

B: City of Johannesburg Metropolitan Municipality

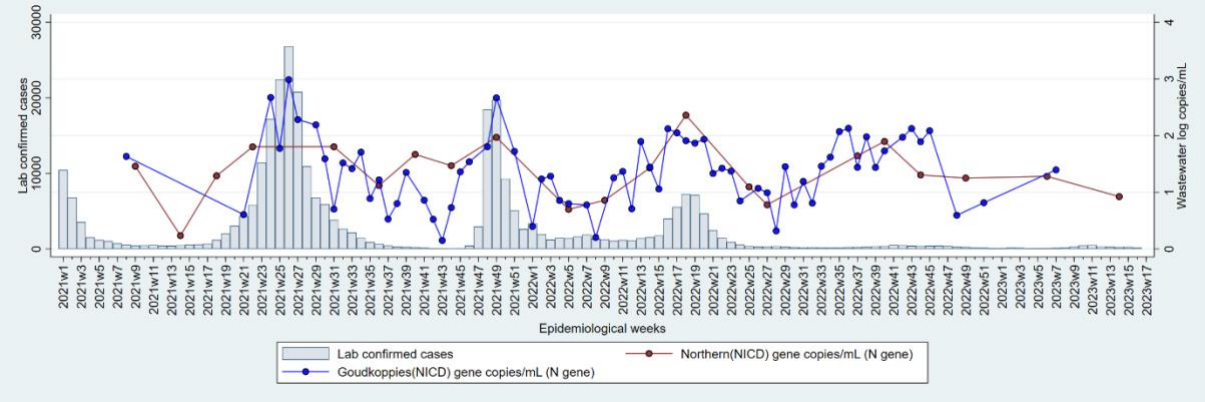
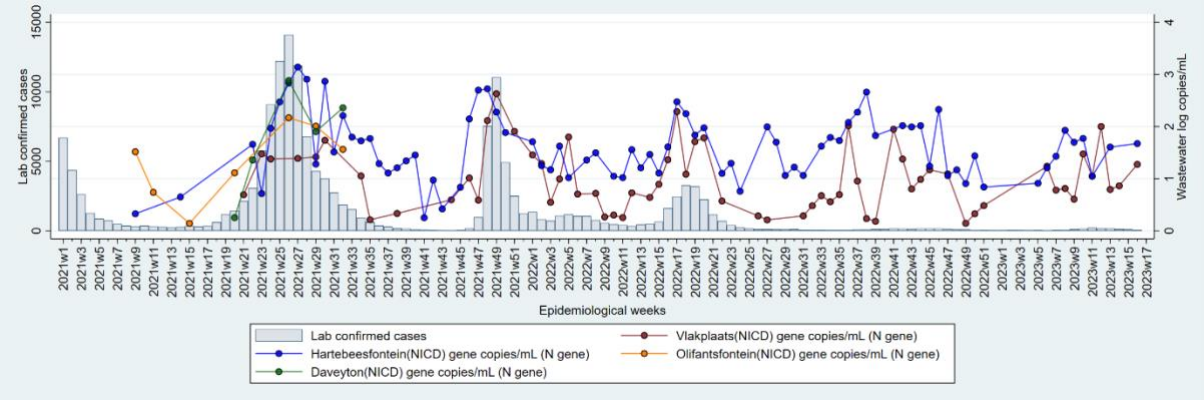


Figure 1B. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) for selected wastewater treatment plants (WWTPs) in the City of Johannesburg Metropolitan Municipality, Gauteng Province during epidemiological weeks 1 of 2021 to week 16 of 2023.

C: **City of Ekurhuleni**



Figures 1C. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) for selected wastewater treatment plants (WWTP) in Ekurhuleni Metropolitan Municipality, Gauteng Province during epidemiological weeks 1 of 2021 to week 16 of 2023.

In epi week 16, moderate levels of SARS-CoV-2 were detected at the Daspoort and Rooiwal WWTPs in Tshwane South and North, respectively. Similarly, the SARS-Cov-2 levels in Vlakplaats WWTP and Hartebeesfontein WWTPs in Ekurhuleni are slightly moderate. Low levels were detected in epi week 15 at the Northern WWTP in Johannesburg in epi week 15.

KwaZulu-Natal Province

2: eThekwini Metropolitan Municipality

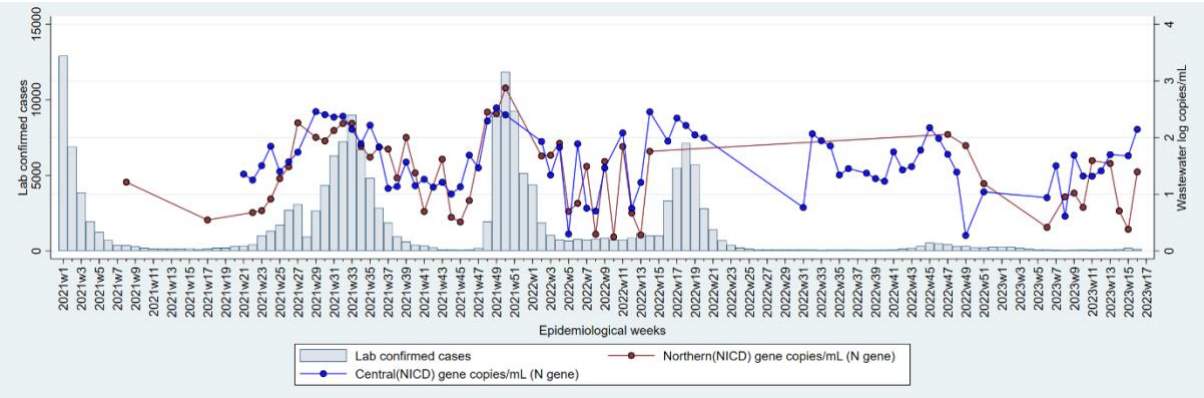
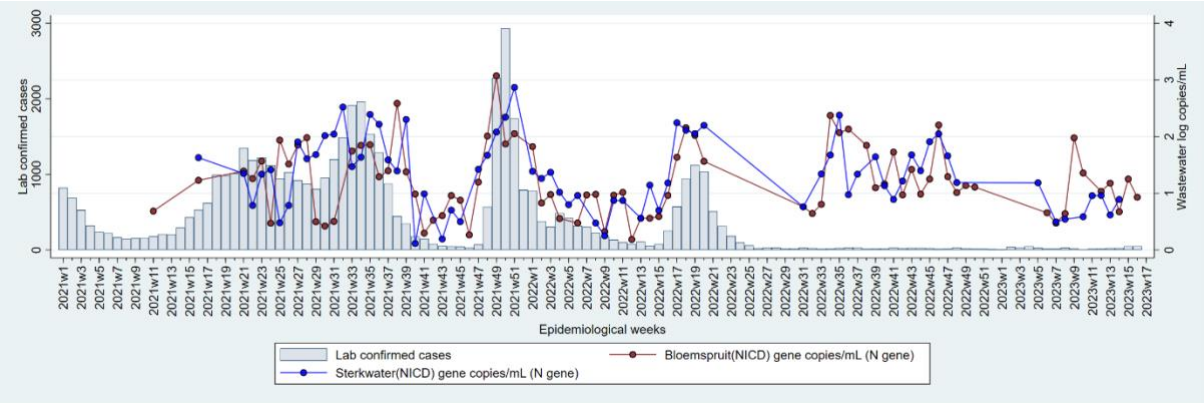


Figure 2. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) from wastewater treatment plants (WWTP) in eThekwini, Kwa-Zulu Natal Province during epidemiological weeks 1, 2021 and week 16, 2023.

Minimal but consistent increases have been observed in the Central WWTP in eThekwini over the last four weeks, with the current level being slightly higher than moderate. In epidemiological week 16, SARS-CoV-2 2023 levels in Central WWTP are moderate.

Free State Province - Mangaung

A: Bloemfontein sub-district



Figures 3. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) from wastewater treatment plants (WWTPs) in Mangaung, Free State Province (Bloemfontein during epidemiological weeks 1, 2021 to 16, 2023.

SARS-CoV-2 levels in Bloemspruit WWTP in Mangaung remained low in epi week 16.

Eastern Cape Province

A: Nelson Mandela Metropolitan Municipality

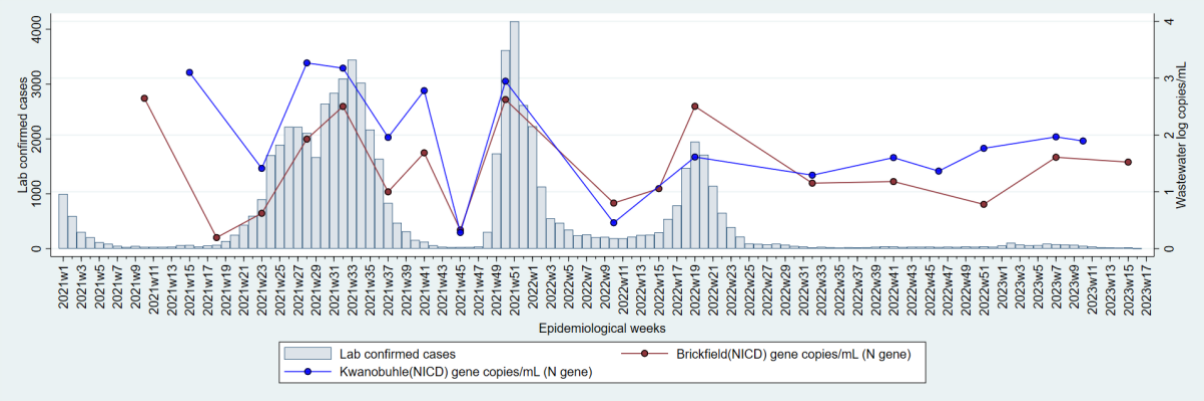


Figure 4A. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) from wastewater treatment plants (WWTPs) in Nelson Mandela Metro, Eastern Cape Province during epidemiological weeks 1, 2021 to 16, 2023.

B Buffalo City Metropolitan Municipality

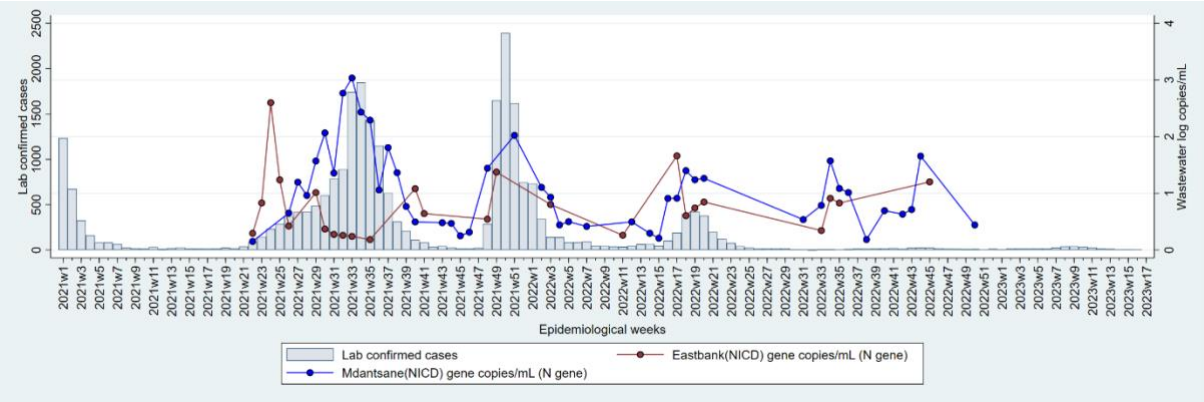


Figure 4B. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) from wastewater treatment plants (WWTPs) in Nelson Mandela Metro, Eastern Cape Province during epidemiological weeks 1, 2021 to 14, 2023.

As of week 16, SARS-CoV-2 levels are at moderate at Brickfield WWTP in Nelson Mandela district. In Buffalo City, the levels at Mdantsane WWTP were low as of epi week 51 2022, requiring latest results. Readers are referred to the SAMRC wastewater dashboard for more in-depth data regarding levels of SARS-CoV-2 in wastewater plants in Nelson Mandela Metro (https://www.samrc.ac.za/wbe/).

Western Cape Province

City of Cape Town

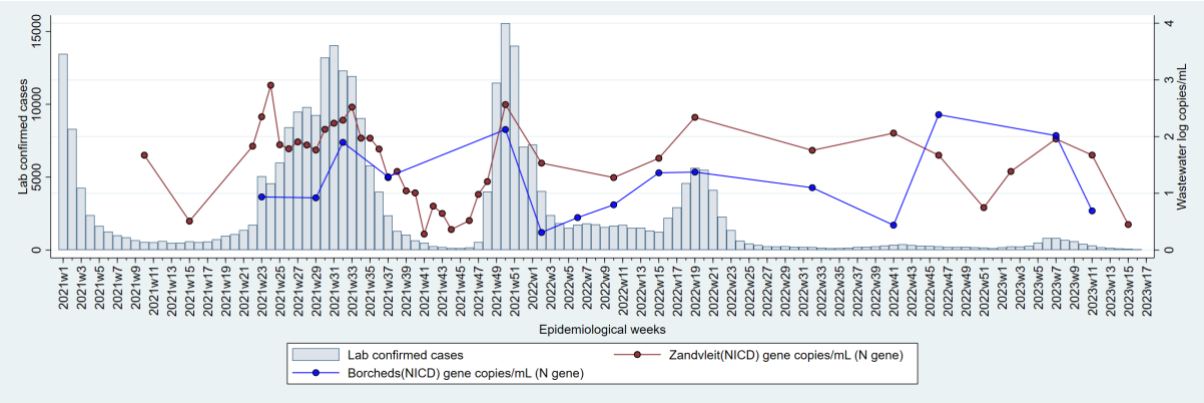


Figure 5. Laboratory confirmed cases of SARS-CoV-2 (bars) and levels of SARS-CoV-2 in log copies/ml of wastewater (coloured lines) from wastewater treatment plants (WWTPs) in the City of Cape Town, Western Cape Province during epidemiological weeks 1, 2021 to 16, 2023.

Low levels of SARS-CoV-2 in wastewater were detected in Zandvleit and Borched's Quarry WWTPs in epi weeks 16 and 11 respectively. Readers are referred to the SAMRC website, which provides data from additional wastewater treatment plants in the City of Cape Town and other Western Cape districts (https://www.samrc.ac.za/wbe/) to contextualise the results.

Limitations

It is not possible to estimate population burden of disease using wastewater testing of SARS-CoV-2 as sources of variability are multiple, including variation in length and concentration of SARS-CoV-2 excretion by infected persons, variation in degradation rate of viral RNA in wastewater and sampling error. Interpretation of results from the levels of SARS-CoV-2 in wastewater is enhanced when the population served by the wastewater treatment plants is well characterised in terms of SARS-CoV-2 testing rates, health seeking behaviour, hospital admissions and deaths due to SARS-CoV-2, as well as other general indicators of health. Further exploration of the relationship between levels of SARS-CoV-2, local trends in clinical case burden, environmental factors, and test methodology will support the interpretation of observed fluctuations in RNA levels. Quality assessment and inter-laboratory comparisons are underway to ensure participating laboratories are providing consistent and comparable results.

PART 2: Results from sequencing of SARS-CoV-2 RNA fragments in wastewater

Background

SARS-CoV-2 has been classified into different variants, that are continually emerging as a result of viral evolution. These variants acquire or lose mutations coding for various epitopes found on key viral proteins which lead to changes in transmissibility dynamics, response to treatment or ability to evade neutralisation by antibodies. WHO classified SARS-CoV-2 variants into variants of concerns (VOCs) and variants of interest (VOIs). VOCs have included Alpha, Beta, Delta, and Gamma, and Omicron. Of these, Beta and Omicron were first reported in South Africa. VOIs include Lambda and Mu (https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/).

The Network for Genomics Surveillance of South Africa (NGS-SA) monitors the epidemiology of SARS-CoV-2 variants in PCR-confirmed cases in South Africa. In clinical cases, variant detection is performed using whole genome sequencing and other methods such as real-time PCR. During the first wave (June to August 2020), the Wuhan SARS-CoV-2 strain dominated amongst clinical cases while in the second wave (November 2020 to February 2021), the Beta variant was discovered and was predominant. The third wave (May to September 2021) was characterized by the dominance of the Delta variant and the fourth wave (November 2021 to January 2022) by the Omicron variant.

Several groups have sequenced SARS-CoV-2 from wastewater including groups in the Netherlands which generated near whole genome sequence from wastewater (Lara *et al*., 2020). In the United States, wastewater sequencing provided comparable results to clinical testing and contained sequences with previously undescribed mutations before they appeared in clinical samples (Crits-Christoph *et al*., 2021).

Here, we report on SARS-CoV-2 sequences and variants of concern present in wastewater samples collected at sentinel wastewater treatment plants in South African urban metros from week 14 in 2021 to week 13 of 2023.

Methods

Wastewater sites

In 2020, the National Institute for Communicable Diseases commenced with sequencing of influent wastewater samples for SARS-CoV-2 RNA from 15 wastewater treatment plants in metropolitan areas, including five in Gauteng Province, four in Eastern Cape province, two in the City of Cape Town (Western Cape Province), two in Mangaung (Free State Province), two in eThekwini (KwaZulu- Natal Province) (Table 1).

Table 1. Characteristics of wastewater treatment facilities and of samples submitted for SARS-CoV-2 sequencing from these sites, 2021-2023

Province	Metro or District	Plant name	Population size served by the facility	Genomic testing			
				Epidemio-logical week when sequencing started in 2021	Number of samples submitted for sequencing	Number of samples with coverage > 50	% of samples with useable quality sequences
Eastern Cape	Buffalo City Metro	East Bank	141000	15	33	11	33,33
		Mdantsane	112900	25	47	20	42,55
	Nelson Mandela Metro	Brickfield	40000	15	17	13	76,47
		KwaNobuhle	100320	15	20	14	70,00
Free State	Mangaung	Sterkwater	200000	16	67	38	56,72
		Bloemspruit	350000	16	70	51	72,86
Gauteng	Ekurhuleni Metro	Daveyton	100000	20	5	0	0,00
		Hartebeesfontain	100000	14	75	50	66,67
		Vlakplaats	200000	21	66	46	69,70
	Johannesburg Metro	Northern	1200000	14	17	11	64,71
		Goudkoppies	500000	21	56	29	51,79
	Tshwane Metro	Rooiwal	unknown	17	82	46	56,10
		Daspoort	unknown	14	79	46	58,23
KwaZulu-Natal	eThekwini Metro	Northern	316425	17	45	24	53,33
		Central	350000	17	67	43	64,18
Western Cape	City of Cape Town Metro	Borcherd's Quarry	380000	15	15	11	73,33
		Zandvliet	460000	15	34	19	55,88
Total					795	472	

Sample collection, RNA extraction, amplification and sequencing

One litre of grab sewage samples were collected and transported at 4°C. Viruses were concentrated from the sample by ultrafiltration (Ikner, Soto-Beltran and Bright, 2011), and RNA was extracted using the QIAamp Viral RNA kit (Qiagen, GmbH, Germany). SARS-CoV-2 was detected by RT-PCR using Allplex™ 2019- nCoV Assay from Seegene kit (Seoul, Korea). RNA was re-extracted from SARS-CoV-2 positive concentrates and subjected to amplicon-based whole genome sequencing using the Sinai protocol with some modifications (Gonzalez-Reiche *et al.*, 2020). Libraries were prepared using the COVIDSeq Kit (Illumina Inc, USA), and sequencing was performed using Illumina COVIDSeq kits as described in (Bhoyar *et al.*, 2021) at the Sequencing Core Facility at the NICD.

Sequence analysis

The ARTIC protocol for sequence analysis (https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html) was used in the Galaxy pipeline for sequence analysis (RC, 2005). Reads were trimmed and filtered according to published criteria (Khailany, Safdar and Ozaslan, 2020). At least 10 reads required at each nucleotide position for downstream analysis. Mutations present at 10% or less were removed from the analysis. Reads were mapped against the reference genome (Wuhan strain/ NC_045512.2) and amino acid variation was analysed. Table 2 illustrates an example of amino acids variation file (https://usegalaxy.eu/).

Table 2: Illustration of amino acids variations. A shows sample ID. B is QC filter, which is quality indicator. C is the number of reads produced for each sample. D is the effect of the mutation detected in the gene. E is the name of the gene where mutation occurred. F is the mutation detected. G is the frequency of the reads in the mutation.

A	B	C	D	E	F	G
Sample	**QC filtre**	**Number of reads**	**Mutation effect**	**Gene**	**Mutation**	**Frequency of mutations**
ENV-COV-21-285_S337_001.fastq	PASS	12	NON_SYNONYMOUS_CODING	ORF1ab	K790Q	0.833333
ENV-COV-21-285_S337_001.fastq	PASS	644	NON_SYNONYMOUS_CODING	ORF1ab	K798N	0.057453
ENV-COV-21-285_S337_001.fastq	PASS	14	NON_SYNONYMOUS_CODING	ORF1ab	F800L	0.857143
ENV-COV-21-285_S337_001.fastq	PASS	44	SYNONYMOUS_CODING	ORF1ab	G45	0.863636
ENV-COV-21-285_S337_001.fastq	min_af_0.05Xmin_dp_1Xmin_dp_alt_10	44	FRAME_SHIFT	ORF1ab	Y46L?	0.045455
ENV-COV-21-285_S337_001.fastq	PASS	1347	NON_SYNONYMOUS_CODING	ORF1ab	T54P	0.123979
ENV-COV-21-285_S337_001.fastq	PASS	153	SYNONYMOUS_CODING	ORF1ab	T54	0.078431

SARS-CoV-2 in the sewage system is fragmented and the genome originated from multiple different individuals, therefore, the generation of a consensus sequence for each sample is not meaningful. Therefore, to identify variants at each geographic location, we analysed amino acid variation in each individual sample. For each VOC or VOI, unique single nucleotide polymorphisms were identified by comparing the new lineage with the Wuhan strain in a public database (https://outbreak.info/). Using the amino acid variation data file, we used STATA software (v 17.1) (https://www.stata.com/) to collate spike-gene mutations in a matrix such that the columns represented the amino acid positions of the spike protein and each row recorded all mutations detected in each sample at every locus across the spike gene, for each epidemiological week. The matrix was used to create a heatmap using conditional formatting on Excel and both low and high read frequency mutations were included. The matrix was also used to plot a mutational profile by filtering out positions were mutations were not present in that respective week and the list of signature mutations present for each VOC and VOI in the spike protein region, listed by WHO (Table 3) were used to deduced the variant or lineage circulating in each week. To further capture evolution and spread of the virus, Freyja, a tool used to estimate the relative abundance of virus lineages present in wastewater. Freyja uses a “barcode” library of lineage defining mutations to uniquely define all known SARS-CoV-2 lineages and solves for lineage abundance using a depth-weighted, least absolute deviation regression approach. Freyja is free to use and available at (https://github.com/andersen-lab/Freyja).

Table 3: Signature mutations and lineages of concern or under monitoring listed and identified by The World Health Organization (WHO) (https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/).

Pango lineage	One or more of these mutations in the spike protein	Relationship to circulating VOC lineages
BA.5	S:R346X. S:K444X. S:V445X . S:N450D or S:N460X	BA.5 sublineages (e.g. BF.7. BF.14. BQ.1)
BA.2.75	**BA.2.75:** S:K147E. S:W152R. S:F157L. S:I210V. S:G257S. S:D339H. S:G446S. S:N460K. S:Q493R **BA.2.75.2:** S:R346T. S:F486S. S:D1199N	BA.2 sublineages
BA.4.6	S:R346T. S:N658S	BA.4 sublineage
BJ.1	S:V83A. S:Y144-, S:H146Q. S:Q183E. S:V213E, S:G339H. S:R346T. S:L368I. S:V445P. S:G446S. S:V483A. S:F490V. S:G798D. S:S1003I	BA.2 sublineage (B.1.1.529.2.10.1.1)
XAY	S:R21G. S:W152L. S:F186L. S:T95I. S:F486P. S:P621S. S:A706V. S:T111I	Recombinant (Omicron and delta)
XBA	S:R21G. S:W152L. S:F186L. S:T95I. S:F486P. S:P621S. S:A706V. S:T111I	Recombinant (Omicron and delta)
XBB.1.5	S:V83A, S:Y144-, S:H146Q, S:Q183E, S:V213E, S:R346T, S:L368I, S: F486P, S:F490V, S:M1233V	Recombinant (BA.2 sub-lineages)
XBB.1.9	XBB + V445P XBB + Orf1a: G1819S + Orf1a: T4175I	Recombinant (BA.2.10.1 and BA.2.75)
XBB.1.9.1	XBB + V445P + S:F486P XBB + Orf1a: G1819S + Orf1a: T4175I, Orf9b: I5T	Recombinant (BA.2.10.1 and BA.2.75)

Results and discussion

Detection of SARS-CoV-2 variants and lineages from wastewater samples using Freyja

Up to the **29th March, 2023**, a total of **795** wastewater samples from sites listed in Table 1 underwent RNA extraction, amplification and sequencing. Of these **795** samples, **472 (59.37%)** yielded SARS-CoV-2 RNA sequences that had a coverage >50%, which were considered for the variants and lineages analysis. Overall, the distribution of SARS-CoV-2 variants in South Africa from wastewater has progressed from the predominance of Beta variant in January 2021, to Delta variant (June 2021) to Omicron in early 2022, which continues to circulate to date with Omicron lineages BA.5 (beige) and BA.2.75 (turquoise) with BQ.1 (neon blue) and XBB.1.5 (mustard) dominance are circulating in February and March in South Africa, as of week 13, March, 2023 (Figure 1) with XBB.1.5 being the dominant lineage (purple), followed by XBB.1.9 (green) and XBB.1.9.1 (olive) (Figure 2).

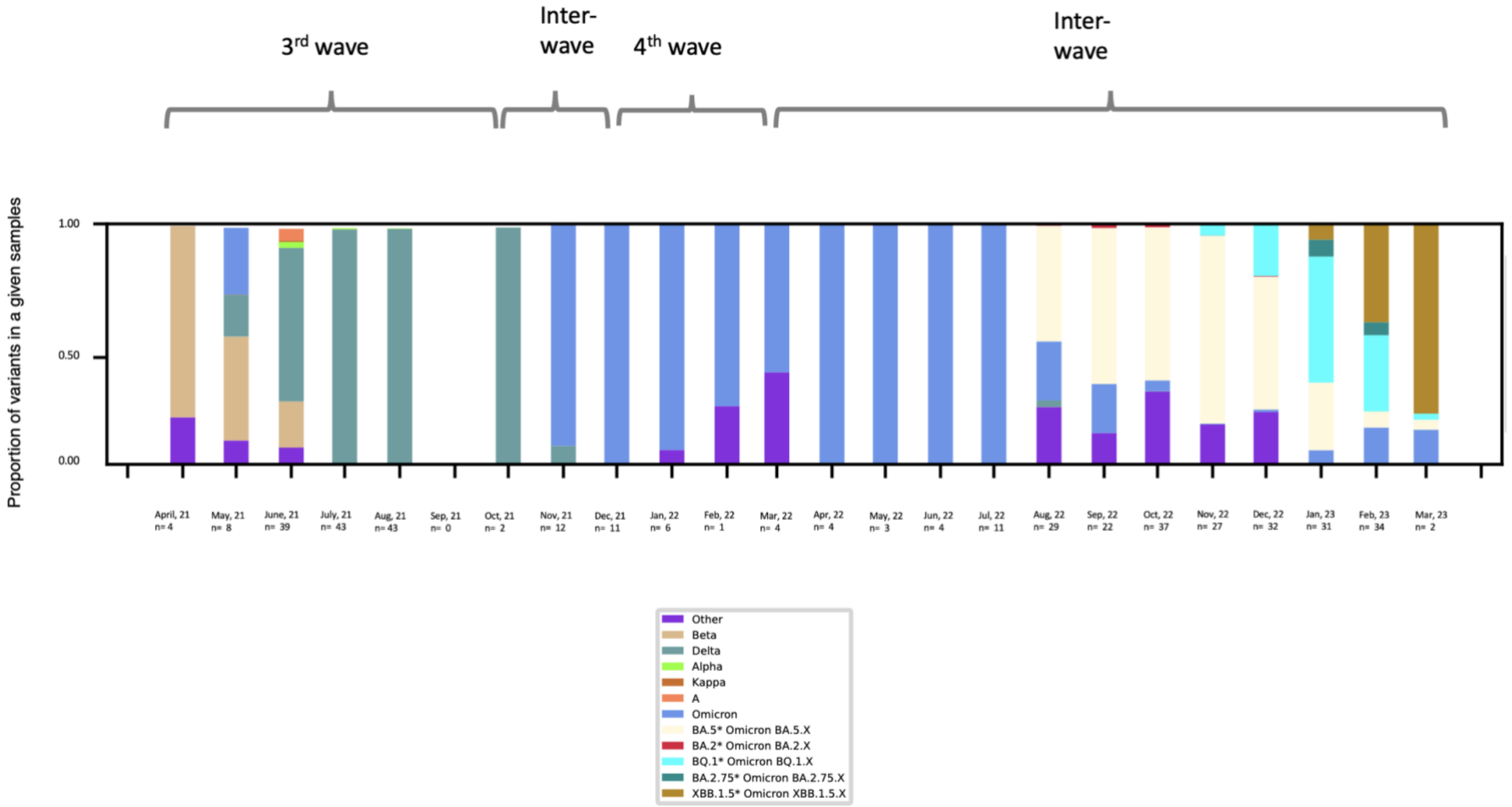


Figure 1. The proportion of SARS-CoV-2 variants in the environmental samples sorted by month and year (April 2021-March, 2023) from all South African provinces. The number of samples processed each month, with a coverage >50% are indicated as n.

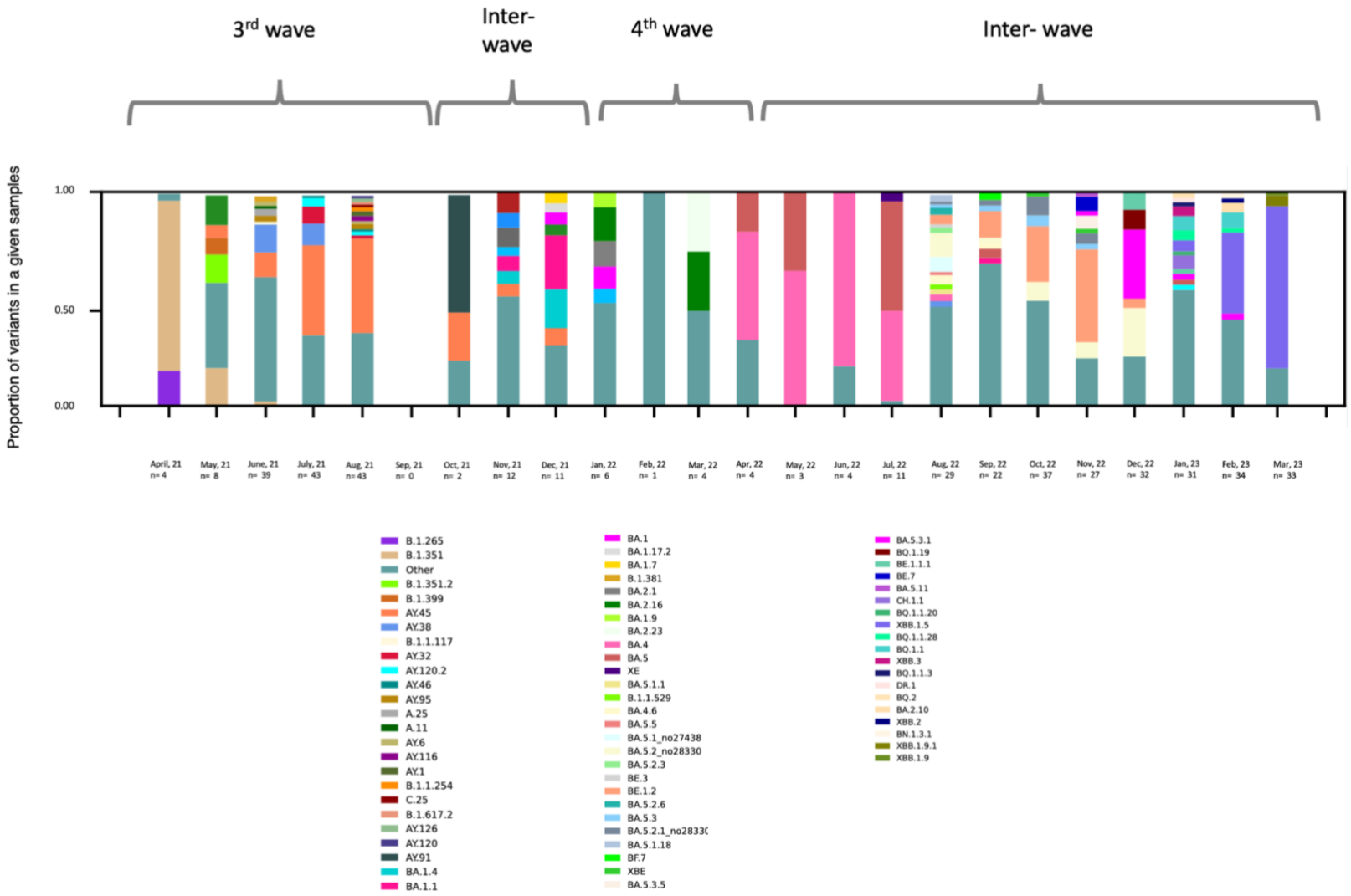


Figure 2. The proportion of SARS-CoV-2 lineages in the environmental samples sorted by month and year (January 2021- March 2023) from all South African provinces. The number of samples processed each month, with a coverage >50% are indicated as n.

Gauteng province

In the Gauteng province**, 228** samples yielded sequencing results displayed in Figure 1-6, which illustrates how Beta variant was present in all the sites in the Gauteng province in week 21-22, 2021 but was replaced by delta shortly after. During the interwave period (weeks 34-44, 2021) most samples submitted for sequencing failed to yield good quality sequence data, most likely due to low or absent SARS-CoV-2 RNA fragments which yielded low coverage . Omicron lineage BA.1 was first detected in week 46, 2021 in wastewater and by week 47, 2021, was found to be present in almost all sites across the province. Lineage BA.2 was then detected from week 5, 2022 in Vlakplaats, followed by the other sites. BA.3 was only detected in Goudkoppies, in week 15, 2022, however at a low read frequency. The low or absence of BA.3 in all other sites was due to either no sampling or low sequence coverage during that period. Gaps in the graph are due to either low coverage or samples were not received during that week. Due to the nature of the wastewater matrix, the genome of certain enveloped RNA viruses like SARS-CoV-2 degrade faster than nonenveloped enteric viruses and therefore have very low coverage. Omicron lineage BA.4 was detected from week 10, 2022 in the Daspoort site and shortly thereafter Omicron BA.5 emerged, causing a resurgence in hospital cases from week 15, 2022. Omicron lineages BE.2, BE.6, BE.7, BE.8, and BE.9 have now since been circulating from week 40, 2022, with BE.1 and BQ.1 dominating in proportion. BQ.1 has also since been detected in clinical case samples, along with lineage XAY (a recombinant lineage between Omicron and Delta, first detected in South Africa) dominating in proportion from week 31, 2022. In wastewater samples however, recombinant XAY was not detected by Freyja in the Gauteng province. In the recent week (week 2, 2023), Freyja has detected BE.1.1, BE.7, BE.8 and BE.9, BQ.1 and 2, BA.5 and XBB.1.5. Lineage XBB.1.5 is a recombinant between BA.2 sub-lineages and was first detected in October, 2022 in the United States of America. Recent sequences from clinical case data in South Africa have also detected XBB.1.5 in 15 patients across the Western Cape, Free State, Gauteng and Kwa-Zulu Natal. XBB.1.5 was detected in wastewater at the Goudkoppies site in week 2. In week 6, 2023, Omicron lineages BQ.1 and BE.1.1 are consistently emerging with recombinant lineage XBB.1.5 showing up in Rooiwal and XBB.3 in Vlakplaats. From week 8 to 13, 2023, recombinant lineage XBB.1.5 has been the dominating circulating lineage in all sites. Lineage XBB.1.9 and XBB.1.9.1 was also observed in the recent week (week 13) to have emerged in Rooiwal, Daspoort, Hartbeesfontein and Vlakplaats, at relatively moderate proportions. XBB.1.9.1 is a new recombinant subvariant that is suddenly emerging across the globe including Indonesia, South East Asia and Europe.

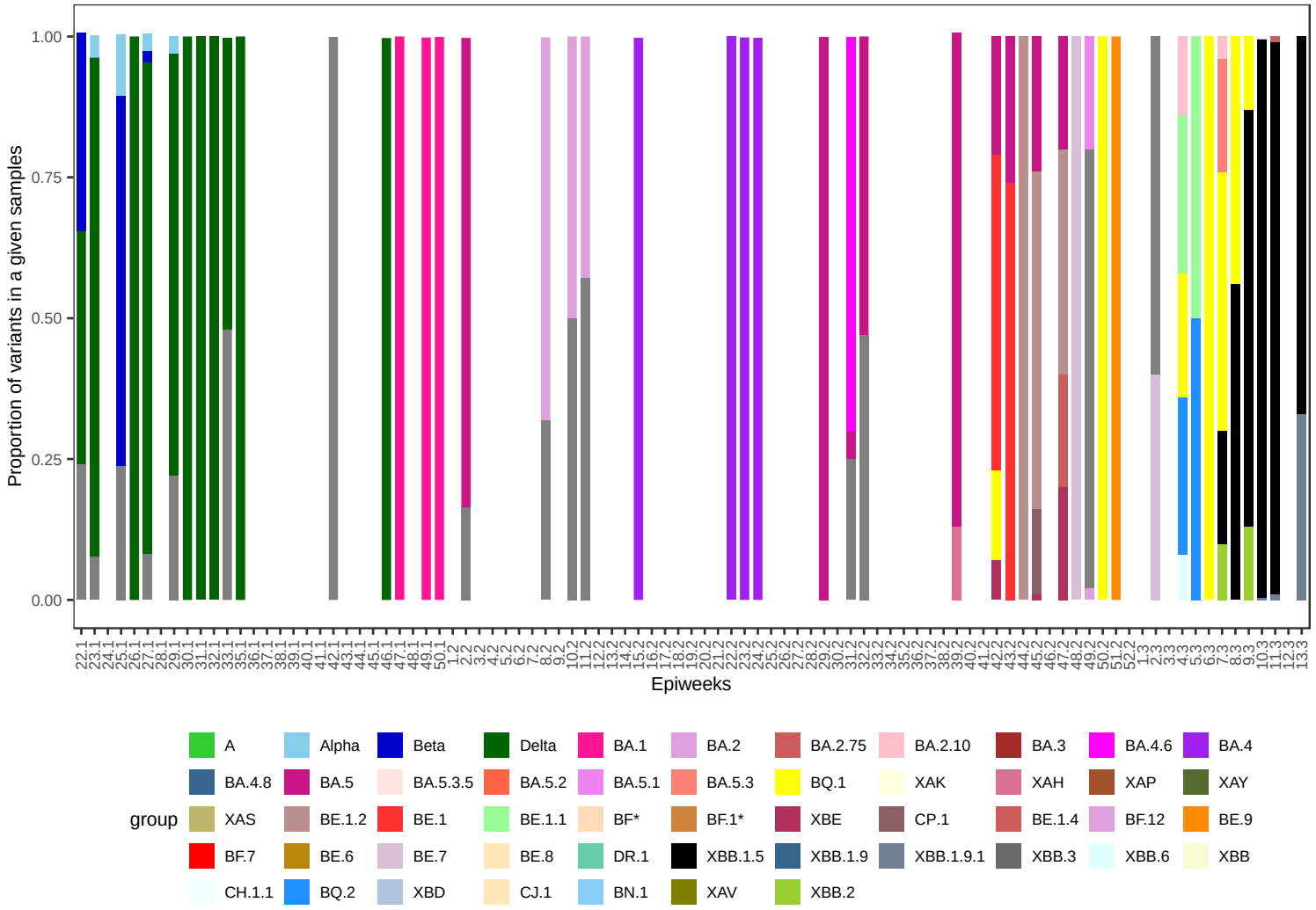


Figure 1: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Hartbeesfontein, in the Ekhurukeni region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

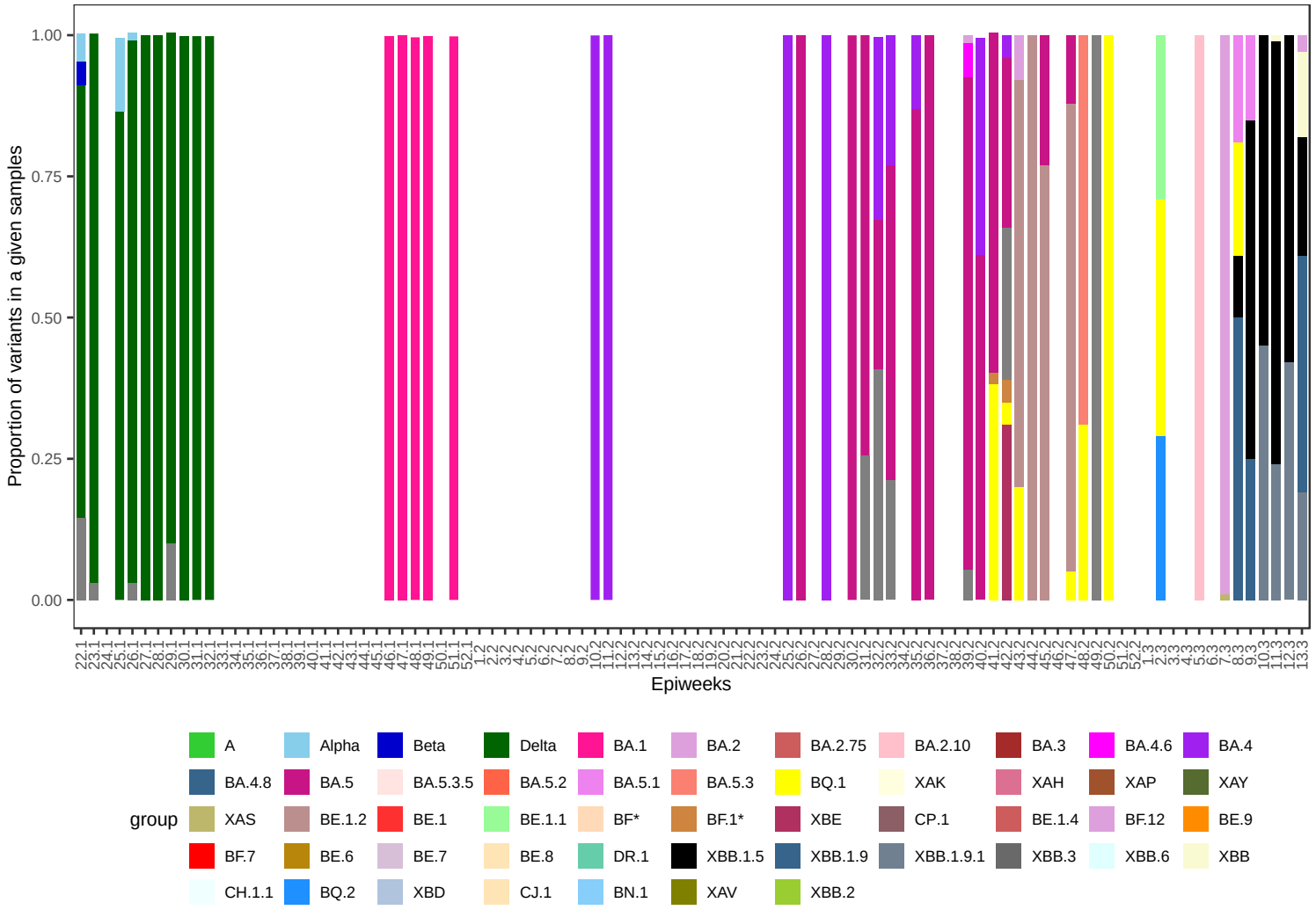


Figure 2: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Daspoort, in the Tshwane region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

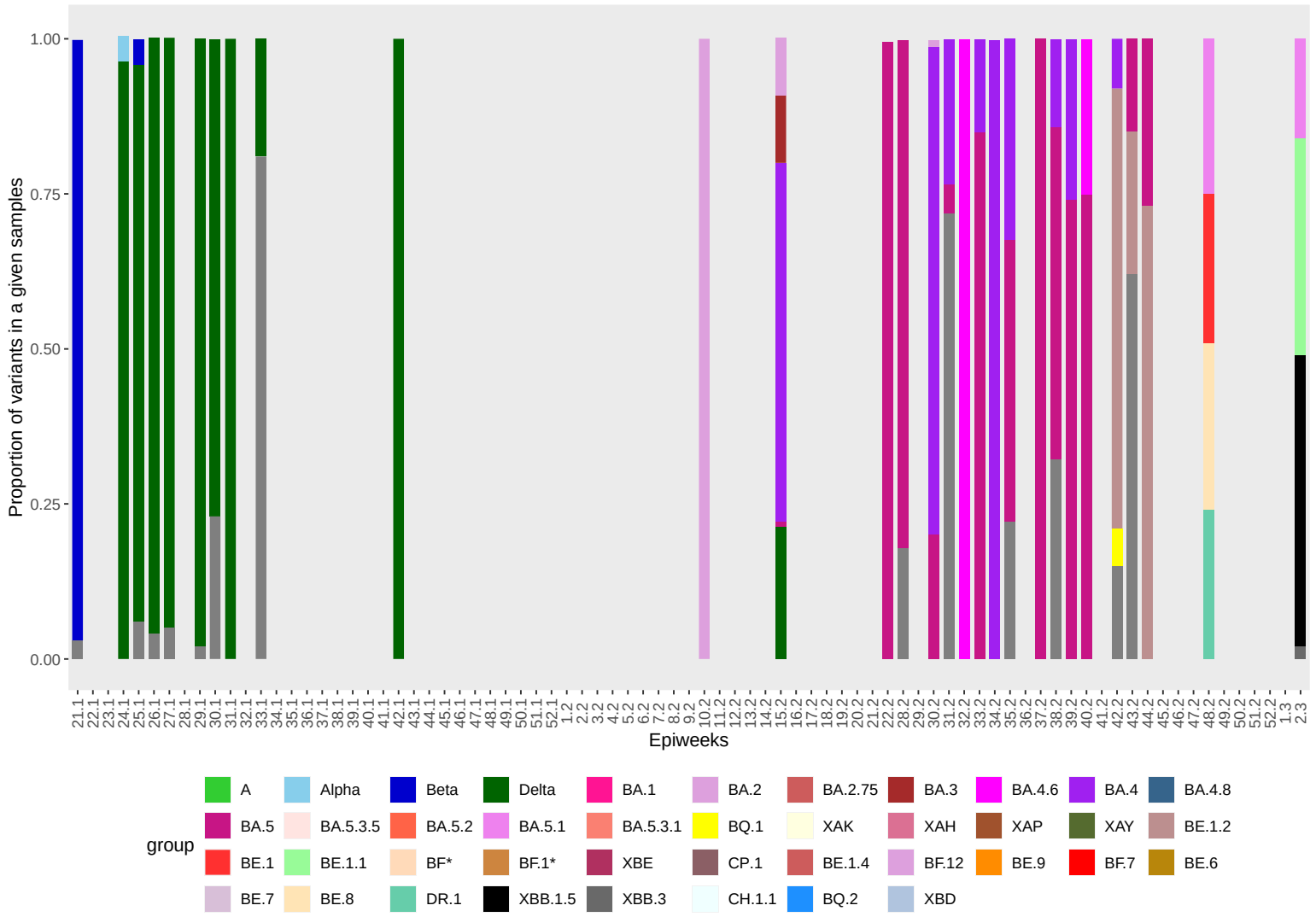


Figure 3: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Goudkoppies, in the Johannesburg region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

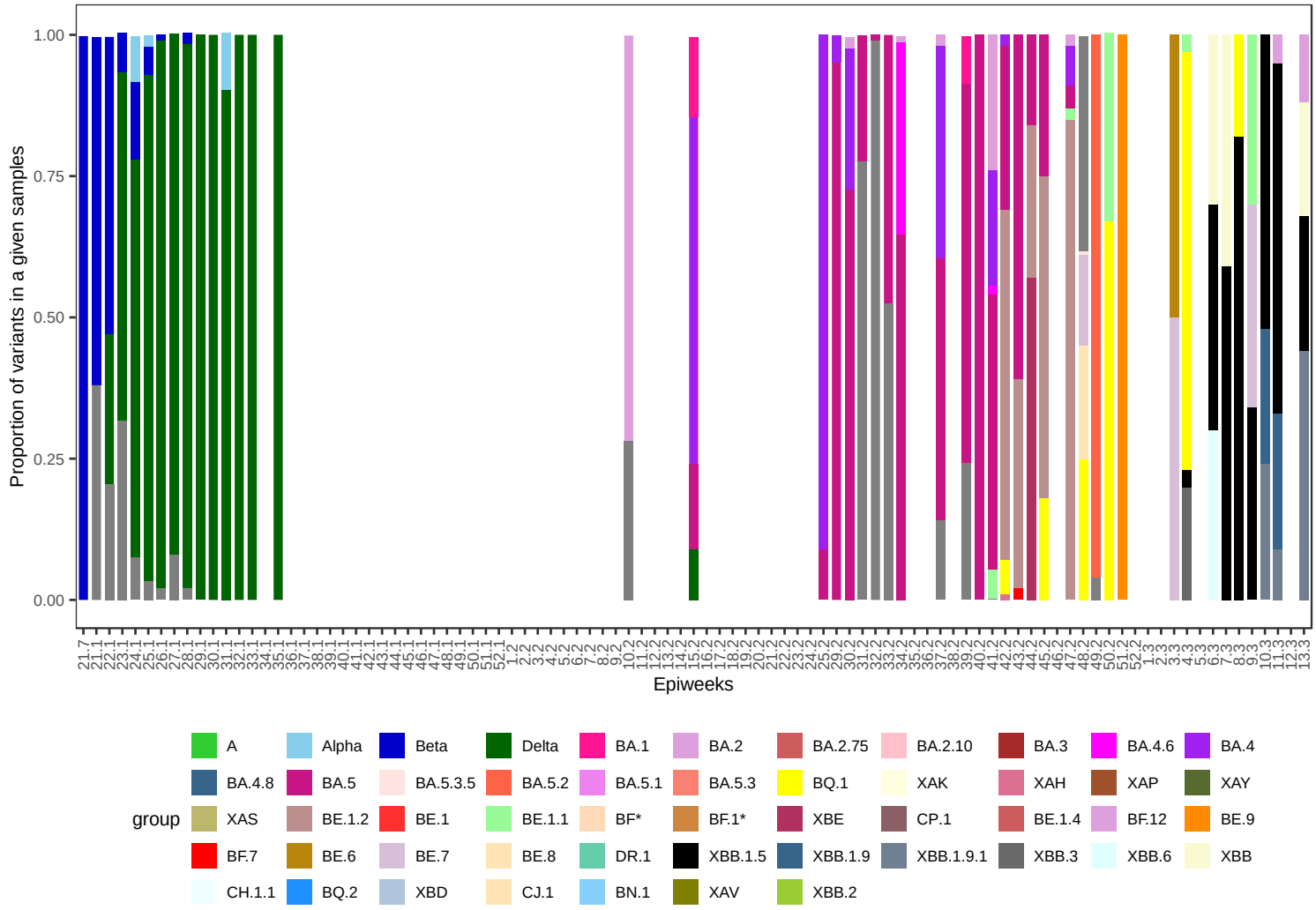


Figure 4: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Rooiwal, in the Tshwane region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

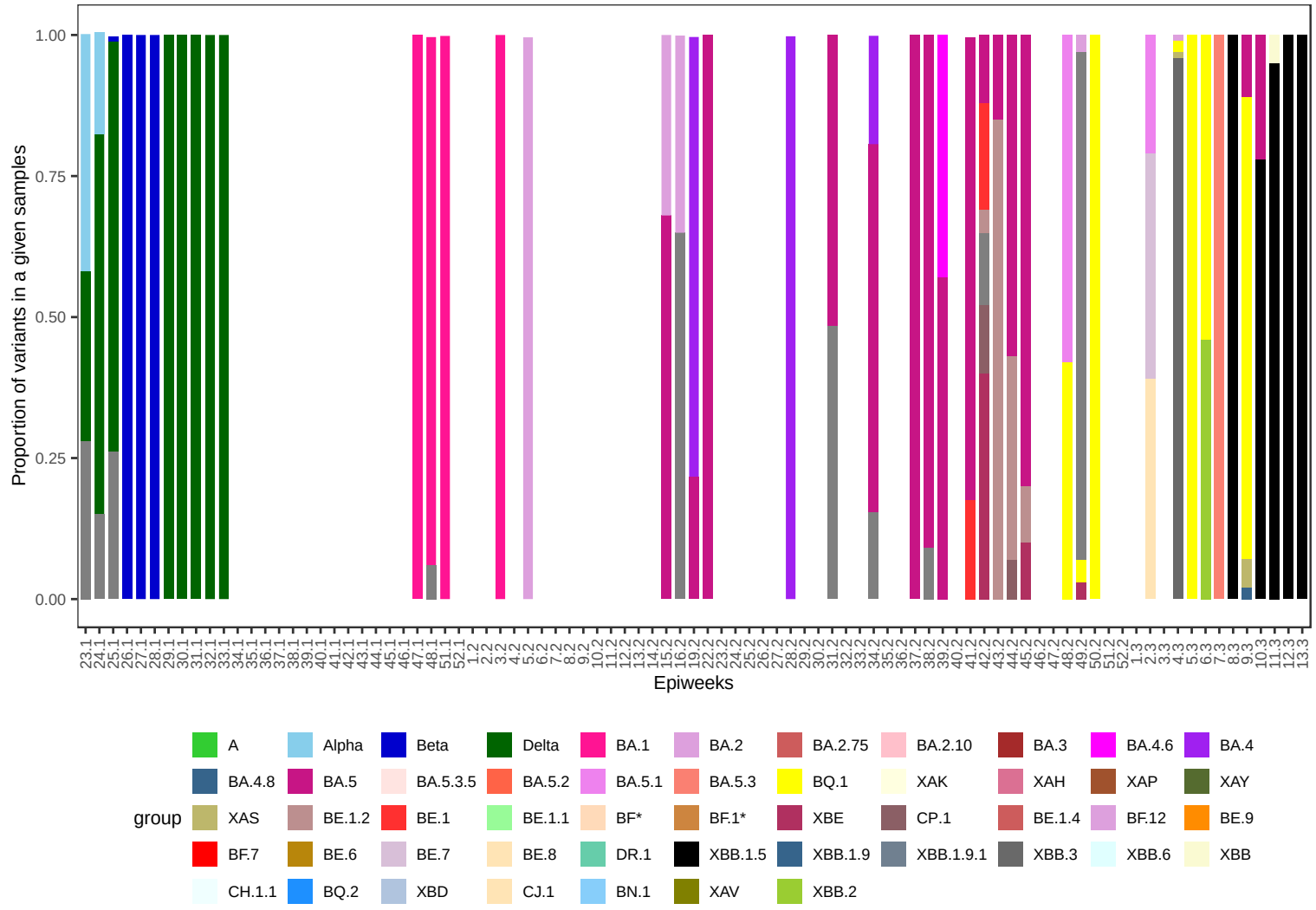


Figure 5: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Vlakplaats, in the Ekhuruleni region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

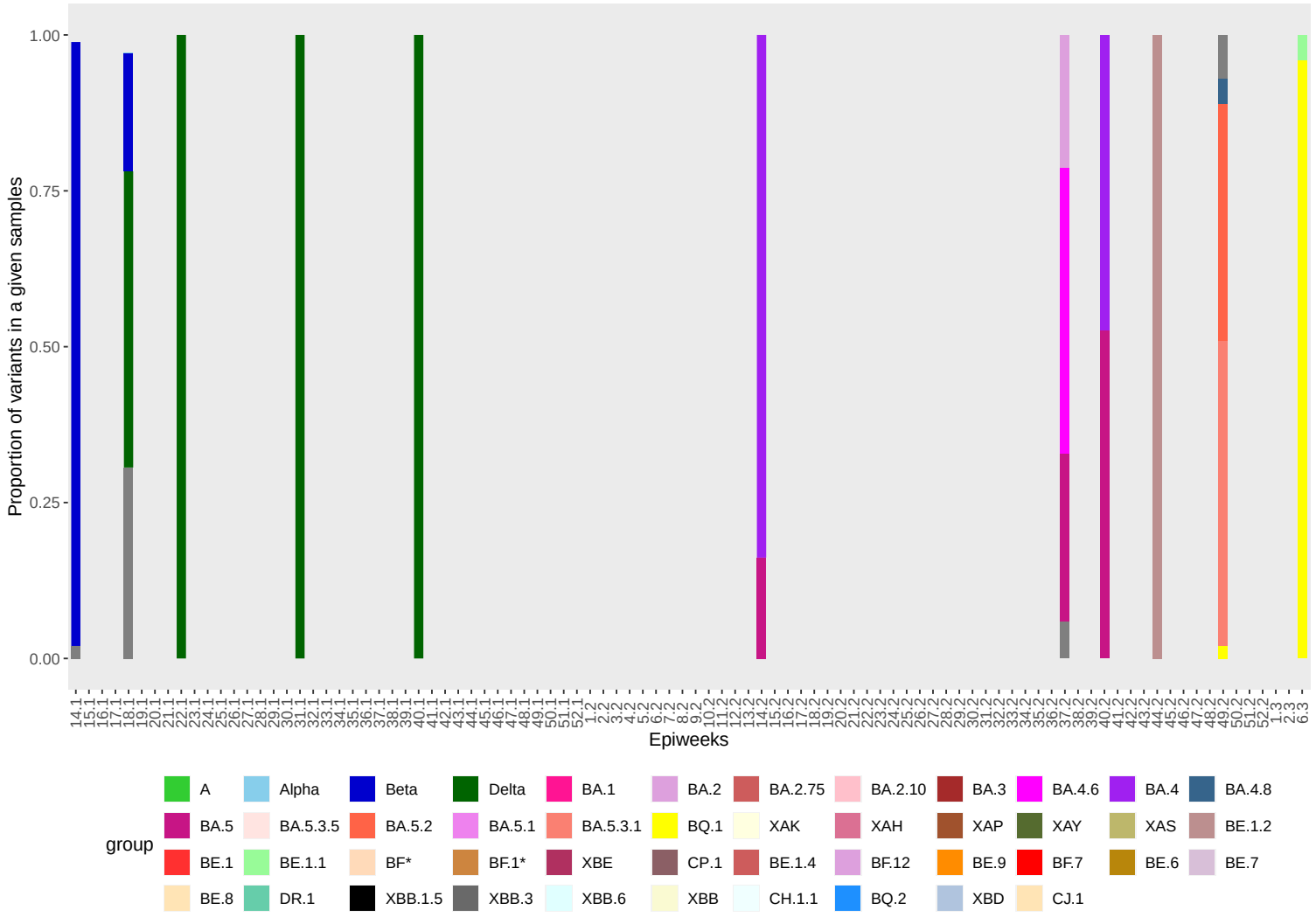


Figure 6: The proportion of SARS-CoV-2 variants and lineages in environmental samples collected from Northern Johannesburg, in the Johannesburg region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

KwaZulu- Natal province

In KwaZulu-Natal province, **67** samples yielded good sequences and were included in the analysis by Frejya. Results are represented in Figure 7 and 8. The Beta variant was detected in a single sample from Central eThekwini plant in week 24, 2021. Subsequently, Delta was first detected after week 22, 2021 in Central eThekwini, followed by Northern eThekwini, in week 24, 2021. As in the Gauteng Province, during the interwave period (weeks 34-44) most samples submitted for sequencing failed to yield good quality sequence data, most likely due to low or absent SARS-CoV-2 RNA fragments. Omicron lineage BA.2 was first detected in wastewater in week 41, 2021 in central eThekwini and week 9, 2022 in central eThekwini and continued to be present up to week 11 of 2022. Omicron lineage BA.4 and BA.4.6 was then detected from week 14, 2022, in both plants. The low levels or absence of BA.2, BA.3 and BA.4 in both sites was due to either no sampling or low sequence coverage in that period of sampling. Gaps in the graph are due to either low coverage or samples were not received during that week. Due to the nature of the wastewater matrix, the genome of certain enveloped RNA viruses like SARS-CoV-2 degrade faster than nonenveloped enteric viruses and therefore have very low coverage. Omicron lineage BA.5 was found to be detected earlier (week 41, 2021) than BA.1, 2 and 3 in central eThekwini. BA.5 then re-emerged in week 9, 2022 and continues to circulate to date. BQ.1 has recently been detected in central eThekwini, along with lineage XAY, at a low proportion in week 48, 2022. From week 2 up until week 5, omicron lineages; BA.2, XBB.3, XAS were detected in both sites. In the recent (week 6), BQ.1 is consistently emerging in both sites with XBB.1.5 showing up for the first time in week 4 and remerging in week 6 in Northern eThewkini. From week 9 to 13, 2023, recombinant lineage XBB.1.5 has been the dominating circulating lineage in both Central and Northern eThekwini sites, with XBB.1.9 and XBB.1.9.1 emerging at relatively moderate proportions in both sites.

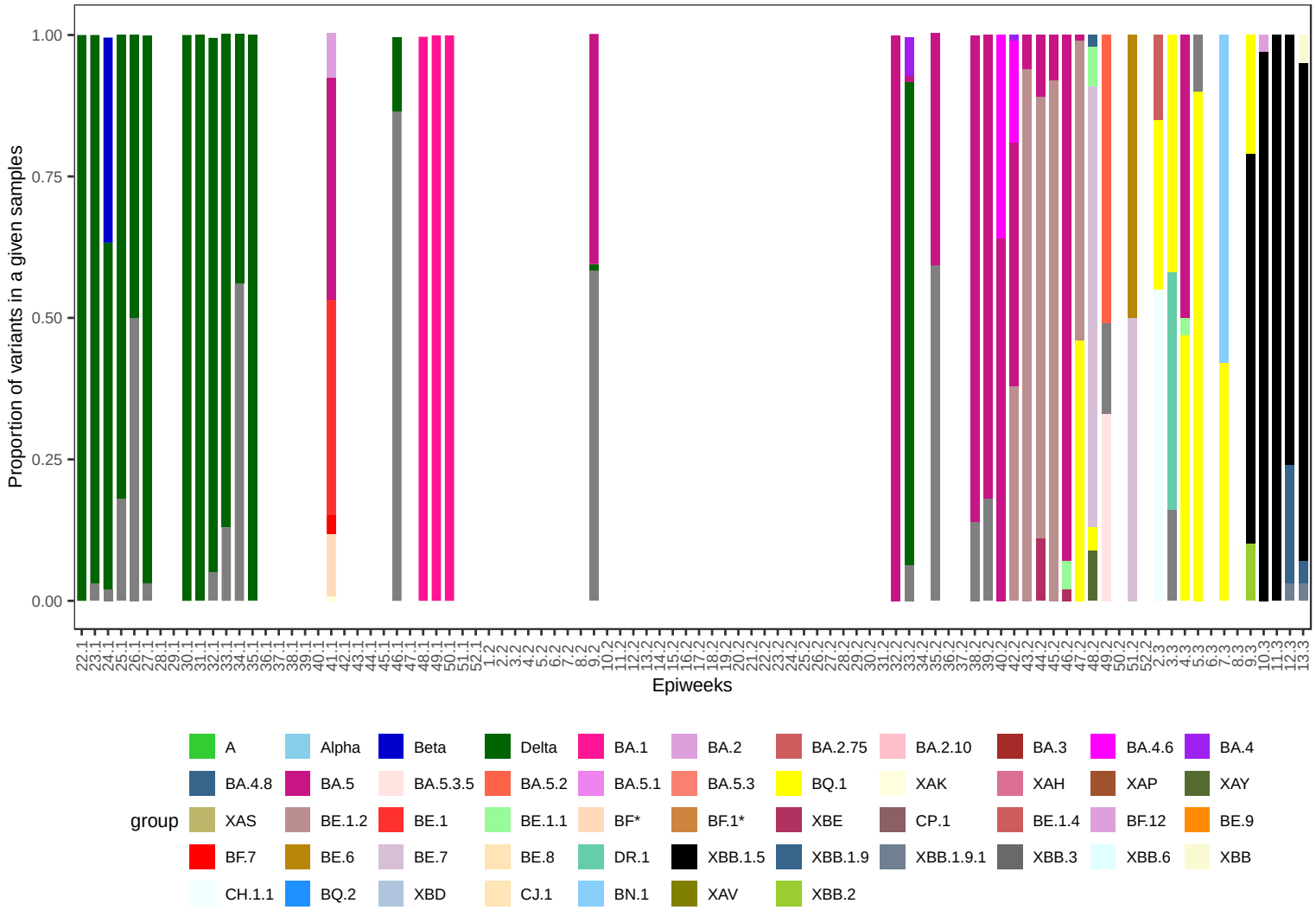


Figure 7: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Central eThewkini, in the eThekwini region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

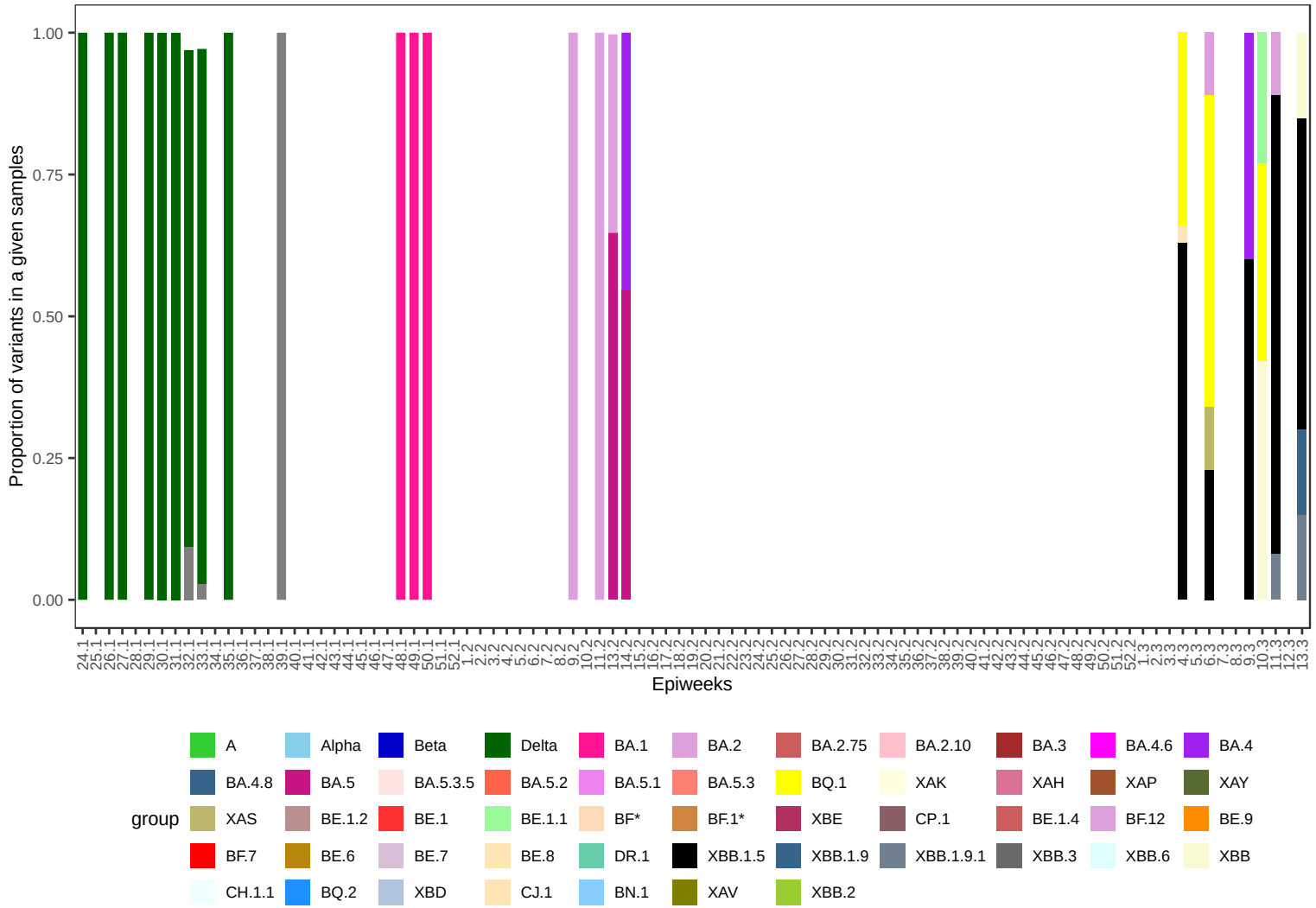


Figure 8: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Northern eThewkini, in the eThekwini region, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

Free State province

In Mangaung, Free State province, **89** samples yielded sequencing results and were analysed by Frejya (Figure 9 and 10). The Beta variant was detected in week 16, 2021 and present until week 25, 2021 in both plants. Alpha variant re-emerged in week 24, 2021 in Bloemspruit. The Beta variant was then replaced by Delta in week 22, 2021 in Sterkwater and week 24, 2021, in Bloemspruit and continued to circulate until week 49, 2021. Gaps in the graph are due to either low coverage or samples were not received during that week. Due to the nature of the wastewater matrix, the genome of certain enveloped RNA viruses like SARS-CoV-2 degrade faster than nonenveloped enteric viruses and therefore have very low coverage. Omicron lineage BA.1 was first detected in week 48 and 51, 2021 at both plants and continued to be present up until week 51, 2021. Lineage BA.2 was detected in week 8, 2022, in Bloemspruit and BA.4 from week 15, 2022 in both plants. Shortly after the emergence of BA.4, BA.5 emerged in both plants after week 31, 2022 and continues to circulate to date, along with BE.1.1 and BE.9 emerging from week 44, 2022. From week 50, 2022 omicron lineage BQ.1 has circulating in Bloemspruit up until week 6, 2023 and in Sterkwater up until week 6, with XBB.3 dominance. From week 8 to 13, 2023, recombinant lineage XBB.1.5 has been the dominating circulating lineage in both sites with XBB.1.9 and XBB.1.9.1 emerging at moderately low proportions.

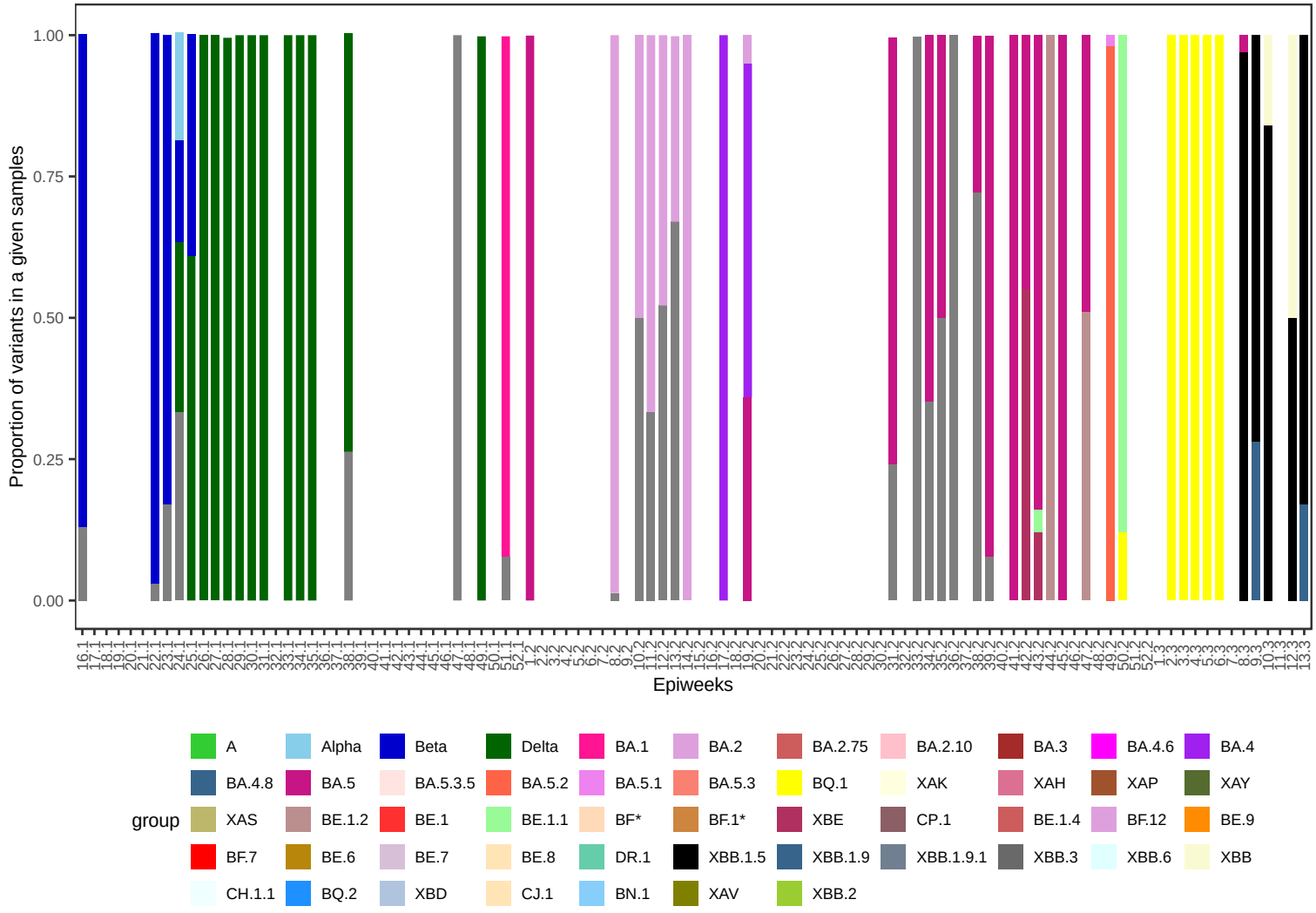


Figure 9: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Bloemspruit, in the Free State, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

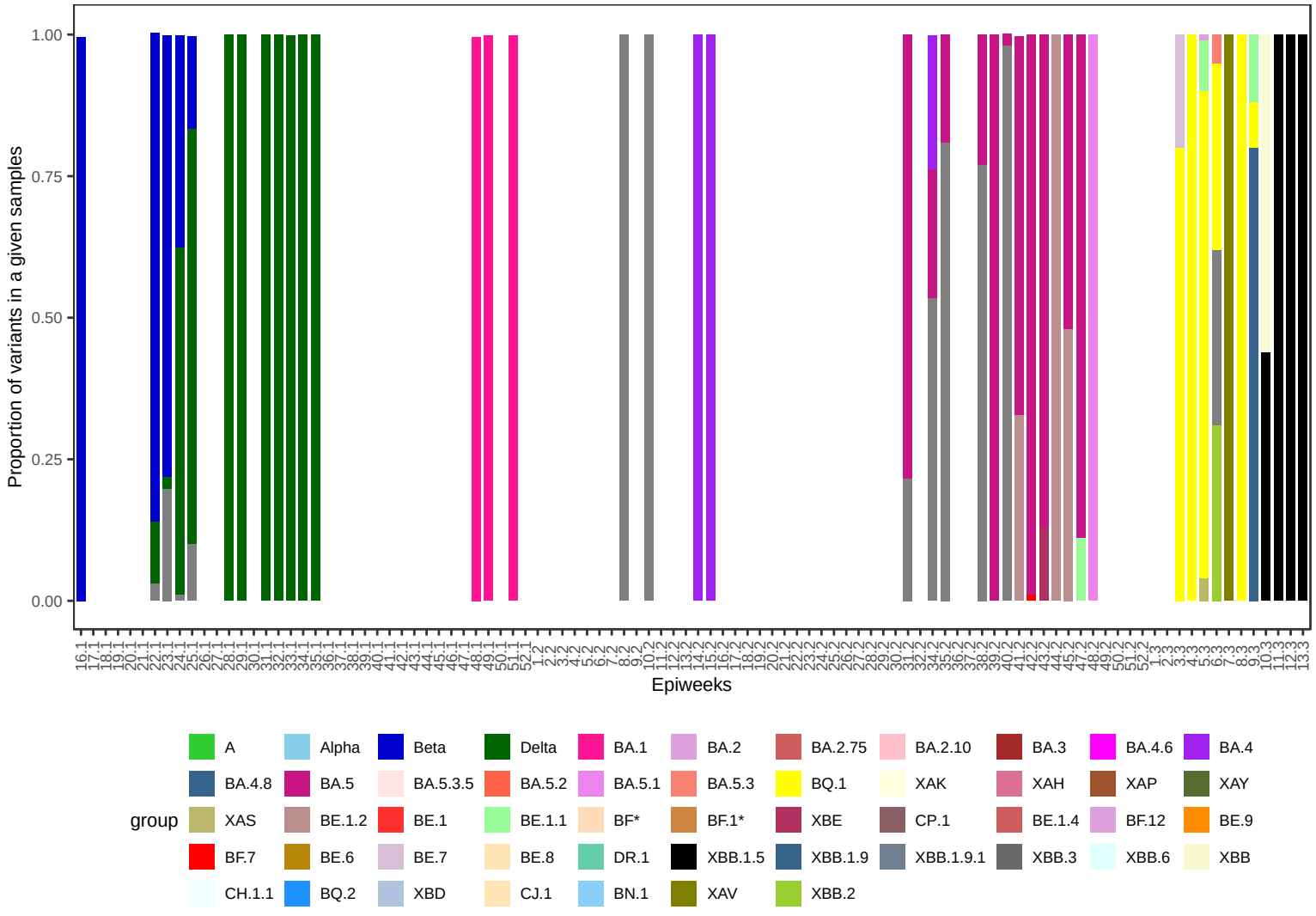


Figure 10: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Sterkwater, in the Free State, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

Western Cape province

In the Western Cape Province, **30** samples yielded sequencing results displayed in Figure 11 and 12. The Beta variant emerged in week 15, 2021 in Borcherd's Quarry and week 22, 2021 in Zandvliet. Beta variant was then replaced by the Delta variant from weeks 23 to 35, 2021. Gaps in the graph are due to either low coverage or samples were not received during that week. Due to the nature of the wastewater matrix, the genome of certain enveloped RNA viruses like SARS-CoV-2 degrade faster than nonenveloped enteric viruses and therefore have very low coverage. Omicron BA.1 was first observed in week 47, 2021 in Zandvliet, followed by BA.2 and BA.3 in week 10, 2022 and week 10, 2022 in Borcherd's Quarry. At both sites, majority of the samples yielded low quality sequence data from week 34, 2021 to week 2, 2022 and week 15, 2022 to week 40, 2022. Omicron lineage BA.5 was first detected in week 41, 2022 and continued to circulate until week 7 with BQ.1and BE.1.1. From week 8, 2023, recombinant lineage XBB.1.5 has been circulating in both sites with XBB.1.9 emerging in Zandvliet, at moderately low proportions.

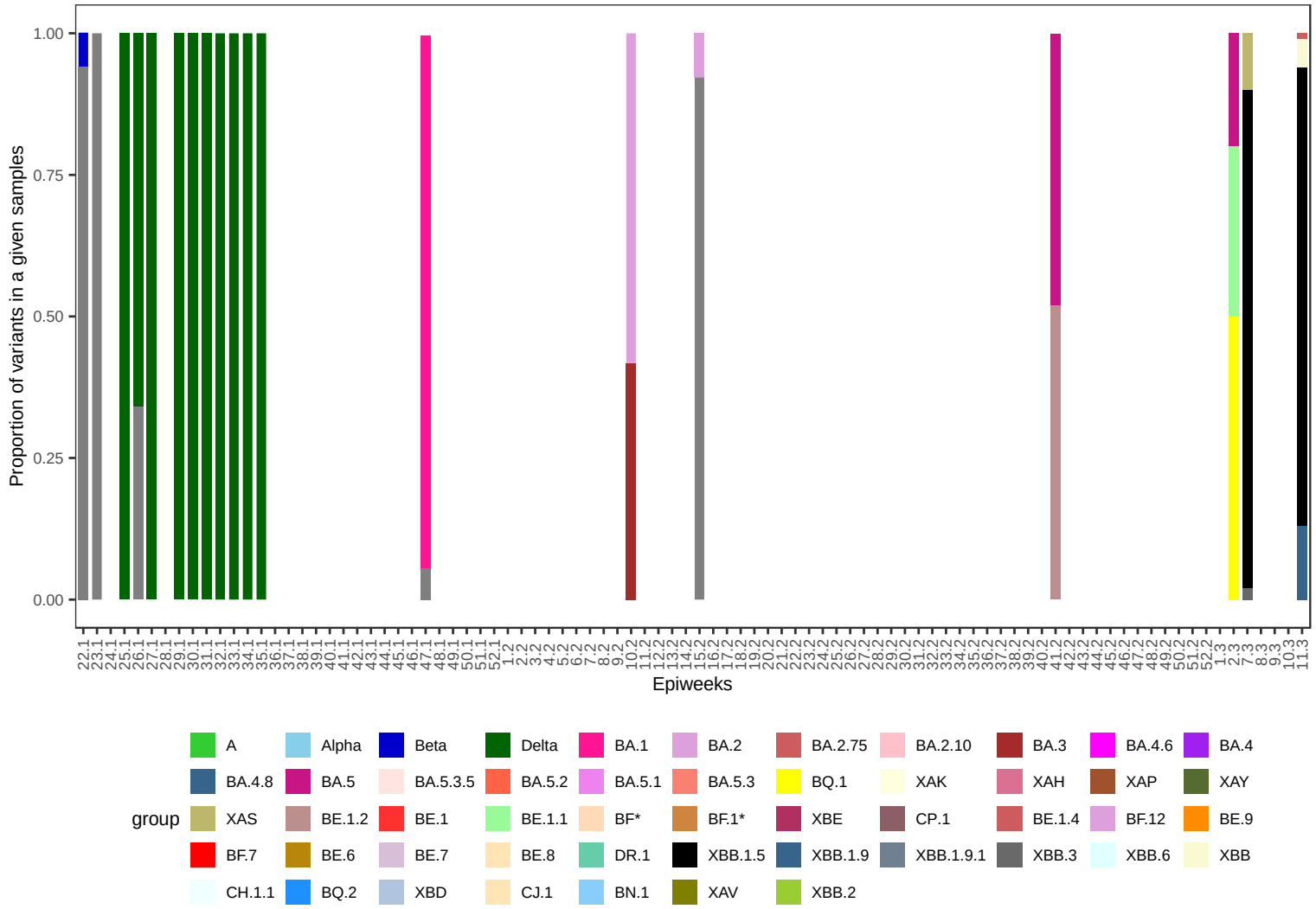


Figure 11: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Zandvliet, in the Western Cape, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

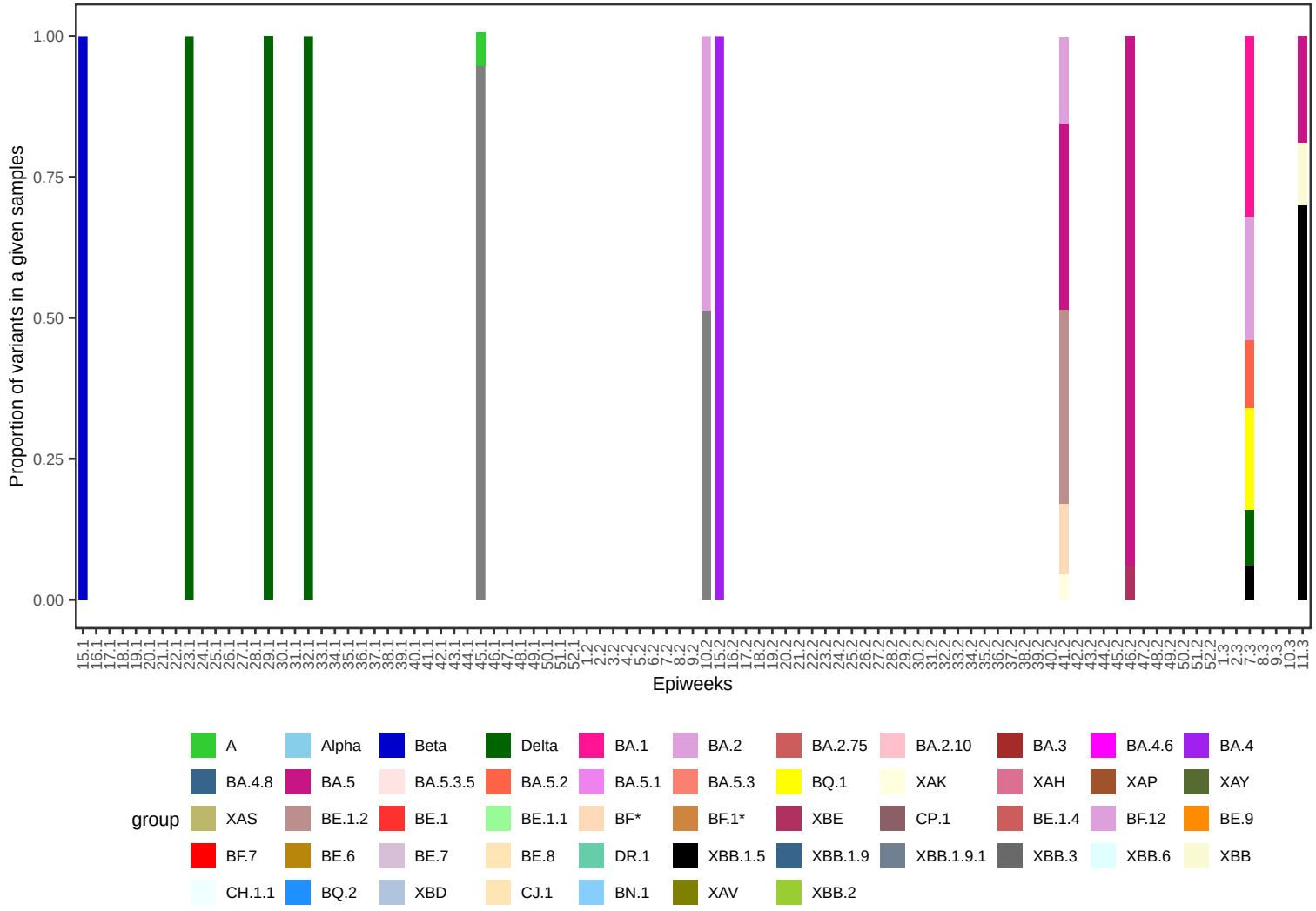


Figure 12: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Borcherd's Quarry, in the Western Cape, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

Eastern Cape province

In the Eastern Cape Province, **58** samples yielded sequencing results displayed in Figures 13,14,15 and 16. The Alpha variant was detected in week 22, 2021 in Eastbank. Delta was first observed in week 23, 2021, in Kwanobuhle and by week 27, 2021, this variant was circulating in all other sites in the Eastern Cape. Gaps in the graph are due to either low coverage or samples were not received during that week. Due to the nature of the wastewater matrix, the genome of certain enveloped RNA viruses like SARS-CoV-2 degrade faster than nonenveloped enteric viruses and therefore have very low coverage. Omicron lineage BA.1 was first detected in week 48 at the Mdantsane site and week 50, 2021 at the Kwanobuhle and Brickfield sites. BA.2 was then detected from week 10, 2022 in Mdantsane, Brickfield and Eastbank. BA.2 was subsequently replaced by BA.4 in week 16, 2022 in all sites except for Brickfield. BA.5 sub-lineages were then detected from week 34, 2022 and continue to circulate in all sites to date, with BE.1, BE.9, CH.1.1 and XBD. From week 8 to 13, 2023, recombinant lineage XBB.1.5 has been the dominating circulating lineage in Kwanobuhle with lineage emerging in Brickfield at moderately low proportions.

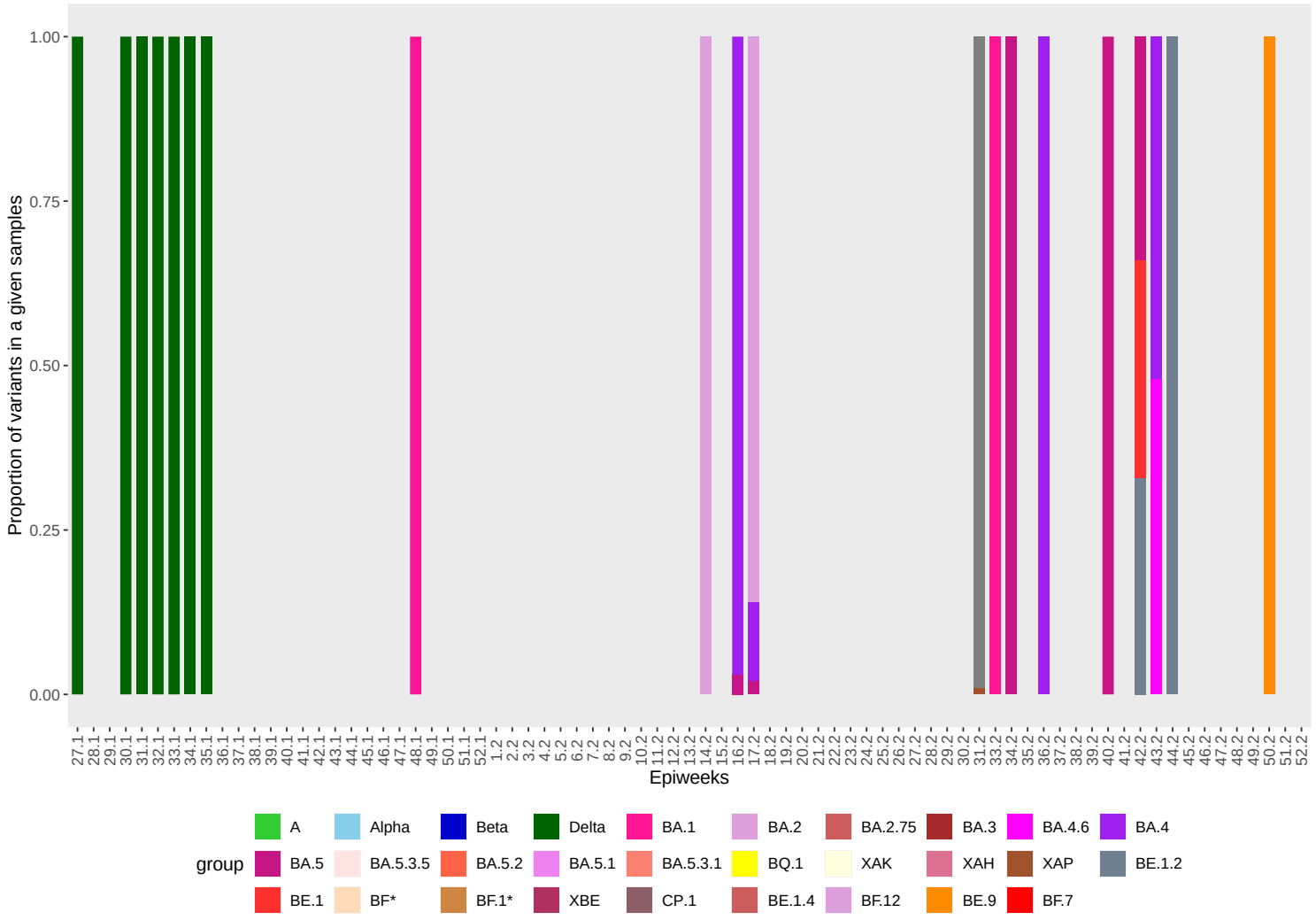


Figure 13: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Mdantsane, in the Eastern Cape, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

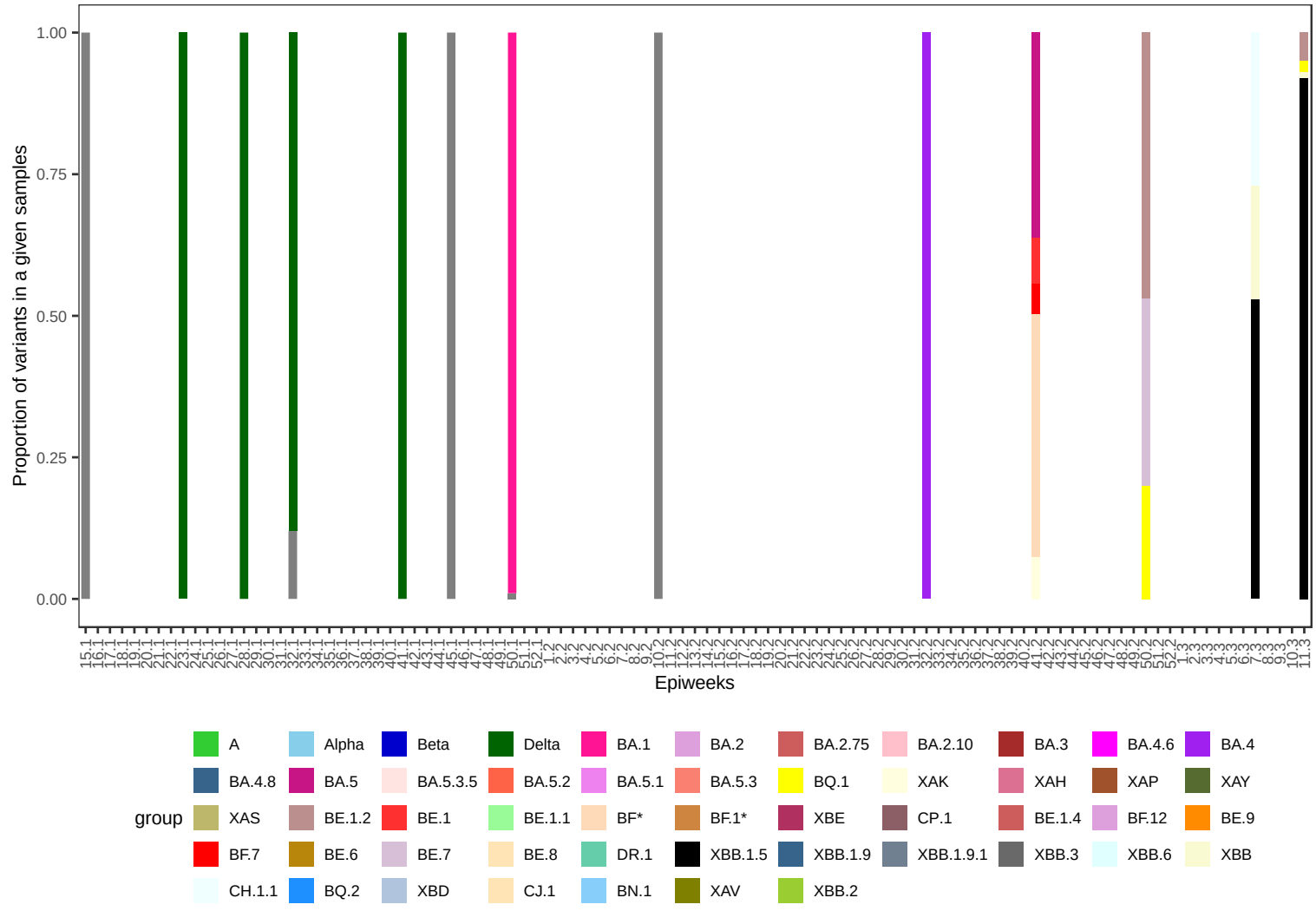


Figure 14: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Kwanobuhle, in the Eastern Cape, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

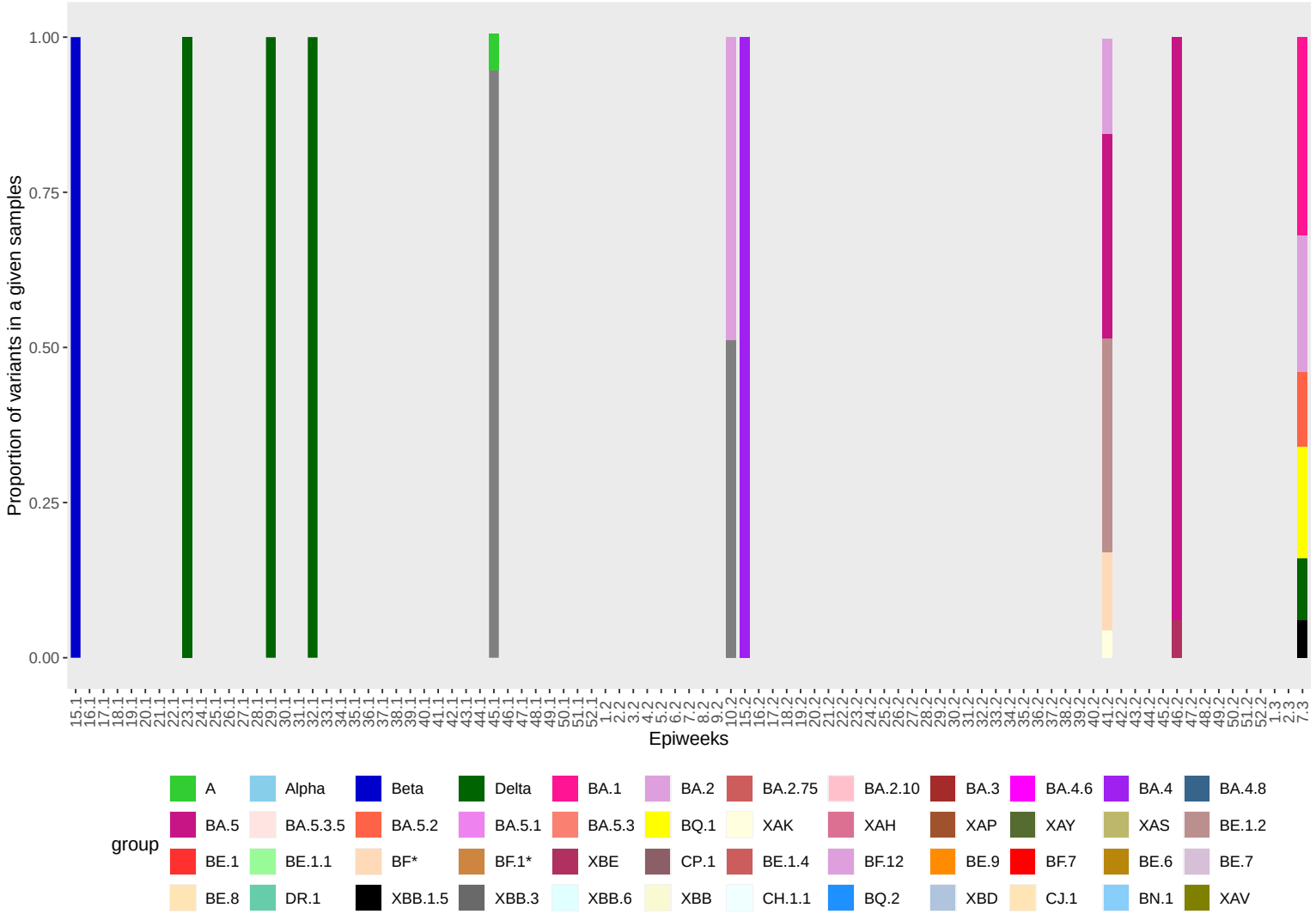


Figure 15: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Brickfield, in the Eastern Cape, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

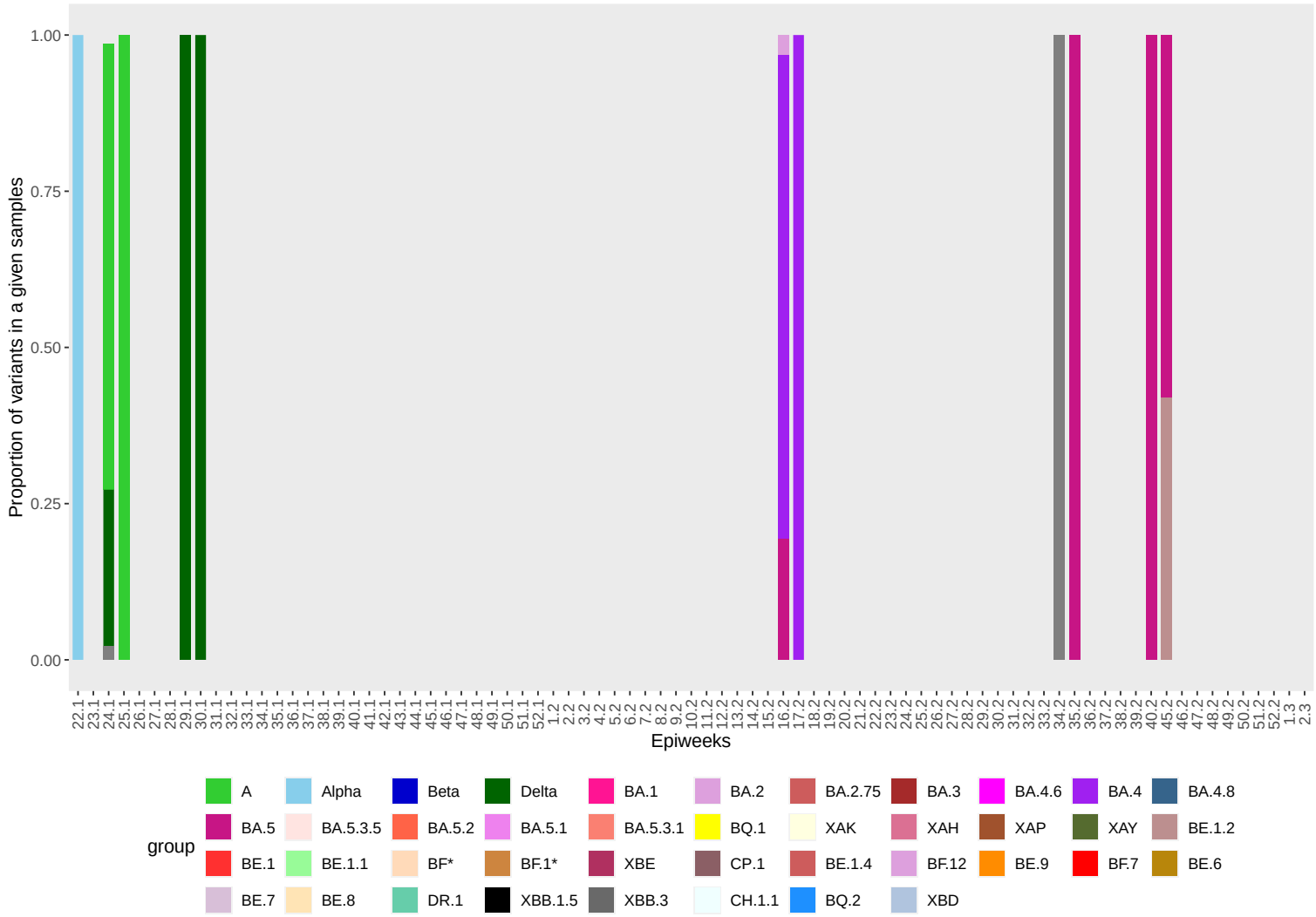


Figure 16: The proportion of SARS-CoV-2 variants and lineages in the environmental samples collected from Eastbank, in the Eastern Cape, arranged chronologically by epidemiological week (i.e. 22.1 is epidemiological 22, year 2021). Only samples that had a coverage of >50% were included in the analysis.

Detection of patterns of emerging SARS-CoV-2 mutations from wastewater samples using a mutational heatmap and mutational profile

A total of **795** wastewater samples from sites listed in Table 1 were used to create a heatmap of patterns of amino acid mutations, starting from epidemiological week 1, 2021 (at the top of the heatmap) to recent week 13, 2023 (at the bottom of the heatmap). In the recent (week 09, 2023), sequencing results and mutations from 7 new samples (from Hartbeesfontein – Gauteng, Vlakplaats – Gauteng, Rooiwal – Gauteng, Daspoort – Gauteng, Central eThekwini – Kwa-Zulu Natal, Northern eThekwini – Kwa-Zulu Natal, Bloemspruit – Free State) have been added the heatmap (Figure 19) and the mutational profile (Figure 20).

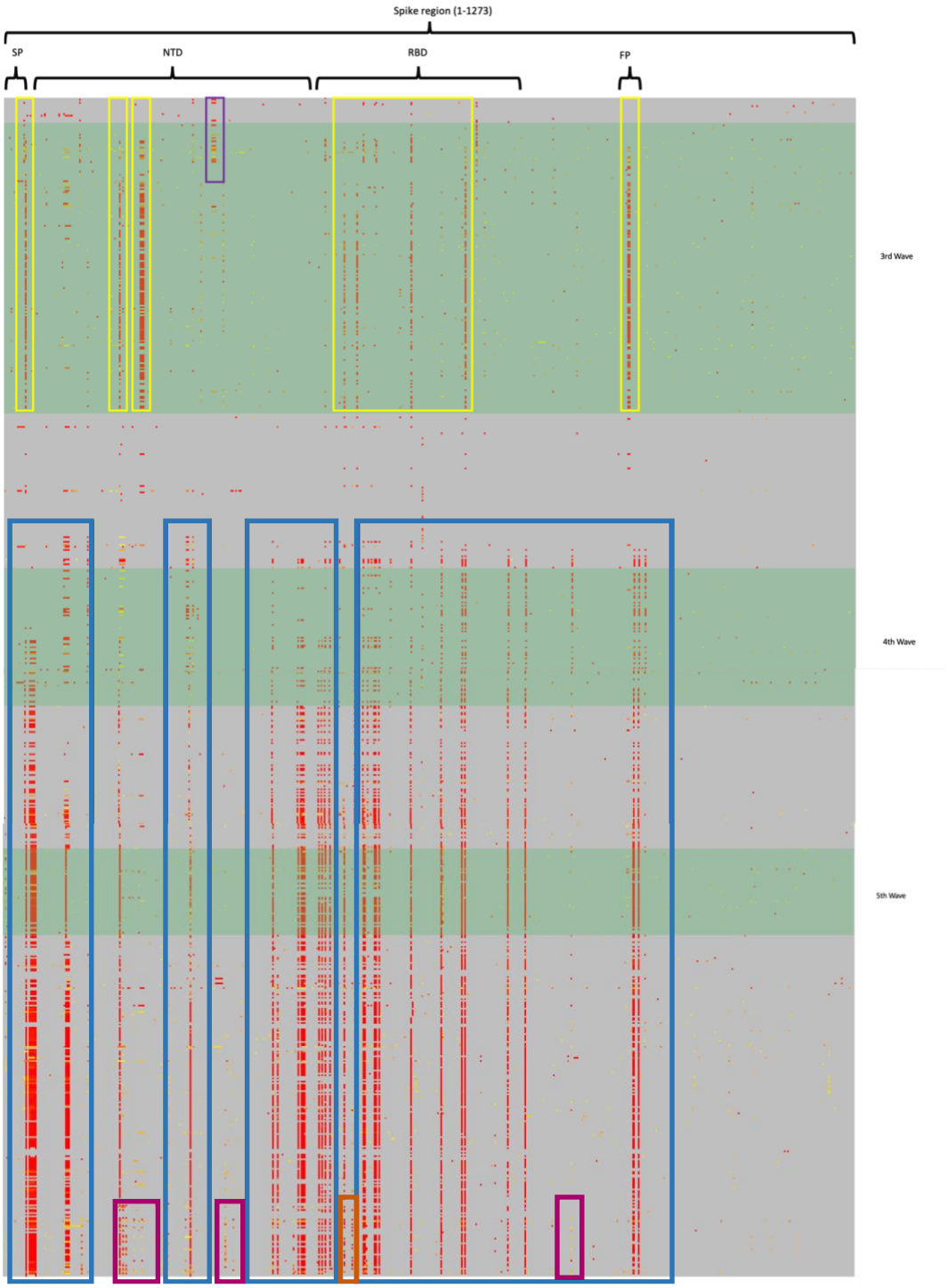


Figure 19: Heatmap of amino acid mutations distributed across the SARS-CoV2 spike protein in comparison with the Wuhan reference strain, arranged vertically in chronological order. Each row represents a sample, organized by the date of sample collection (From April, 2021 to March, 2023). Each column represents an amino acid position of the spike protein. Regions with no mutations or low occurrences are represented in grey (0%) and light yellow (1-34%). Regions with mutations that have a 50% read frequency are represented in dark yellow. Regions with mutations with a read frequency between 60-80% are represented in orange and very high occurring mutations (89-100%) are represented in red. Regions that are highlighted with green represent the time period in which South Africa experienced a wave. Yellow boxes indicate mutations that had emerged that lead to the Delta wave, the blue boxes indicate mutations that had emerged that lead to the Omicron wave and pink bow indicate recent emerging mutations.

The alignment and ordering of the spike amino acid positions in Figure 19 demonstrate characteristic patterns of emerging mutations in epidemiological week. In week 16, 2021, the Delta variant was characterized by the emergence of mutations in the N-terminal domain (NTD) region (G142D, E156del, F157del, and R158G) highlighted in the yellow box, followed by the loss of the N-terminal domain (NTD) region mutations after week 35. This signified the transition from the Delta variant to the Omicron variant. The Omicron variant (highlighted in blue box) was characterized by the emergence of mutations in the receptor binding (RBD) domain (G339D, S371L, 373, N440K, S477N, E484A, Q493R, G496S, Q498R), and fusion peptide (FP) region (N764K, D796Y), and the heptad repeat 1 (HR1) region (Q954H, N969K, L9811F), in week 45, 2021, highlighted in the blue box. Between the third and fourth wave of infection low sequence coverage of spike was observed, likely due to low levels of virus in wastewater because of low clinical caseloads, and few mutations were detected. Mutations (G21R, W152L, F186L, P621S A706V and T1117I) associated with XAY (a lineage first detected in South Africa), were first detected in wastewater in week 20, 2022, and continue to emerge sporadically in specific weeks. From week 48, 2022 up until week 09, 2023 mutations; V83A, Q183E, Y144del, H146Q, W152R, R156del, F157del, R158G, I210V, V213E, L368I, F486P and F490S are consistently re-emerging in the heatmap (highlighted in pink). Mutations; V83A, Q183E, Y144del, H146Q, V213E, R346T, L368I, F486P and F490S mutations are associated with XBB.1.5 and W152R, F157L and I210V are mutations associated with BA.2.75. Mutation V445P associated with XBB.1.9 and XBB.1.9.1 has also been consistently emerging in all samples from week 10 (highlighted in orange).

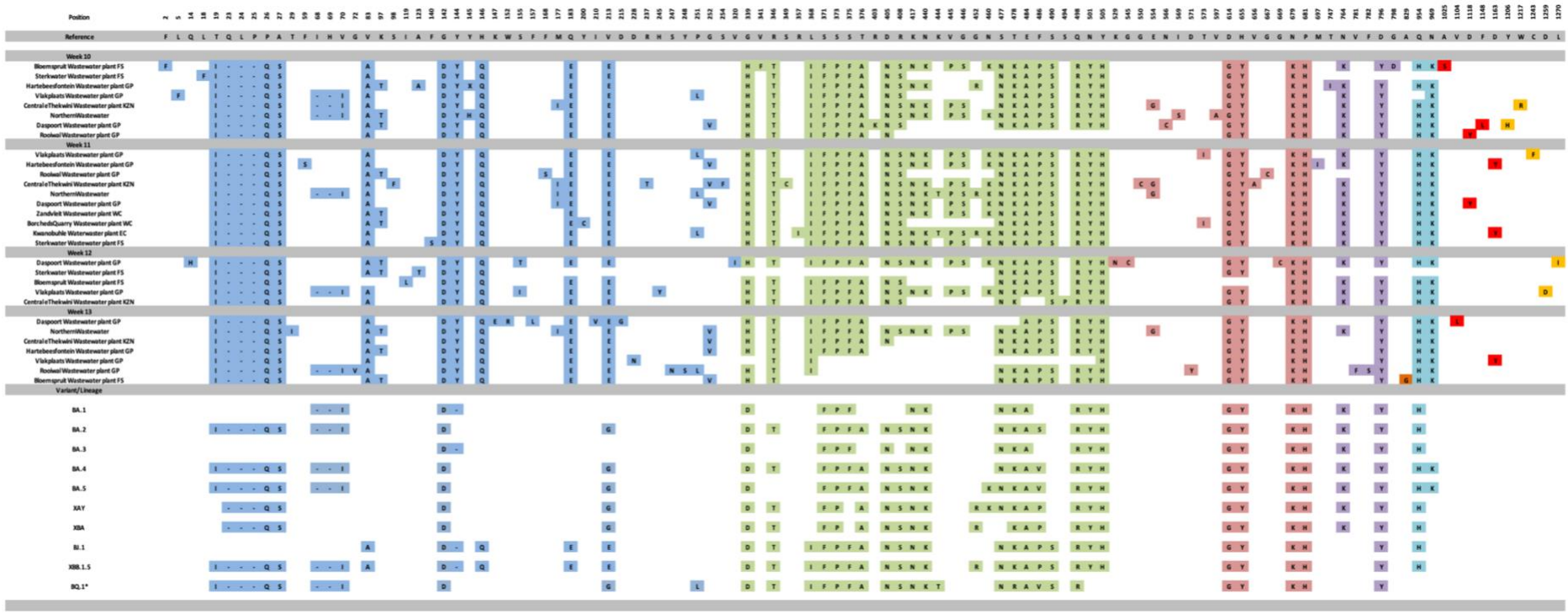


Figure 20: SARS-CoV-2 spike protein mutational profile of samples collected from wastewater sites across South Africa (Zandvliet – Western Cape, Kwanobuhle – Eastern Cape, Brickfield – Eastern Cape, Hartbeesfontein – Gauteng, Vlakplaats – Gauteng, Goudkoppies – Gauteng, Daspoort – Gauteng, central eThekwini – Kwa-Zulu Natal, northern eThekwini – Kwa-Zulu Natal, Bloemspruit – Free State, Sterkwater – Free State) with the respective associated lineage or variant. Each row represents the site in which a sample was collected (top half) and the mutations that are associated with lineages or variants of concern (bottom half). Each column represents an amino acid position of the spike protein, with the wildtype represented below. Mutations are listed within the plot and are colour coded according to the spike region they are found in (NTD – N-terminal domain (blue), RBD – Receptor binding domain (green), SD – Subdomain (pink), UH – Upstream helix (purple), HR1 – Heptad repeat (orange and powder blue), SD3 – Subdomain 3 (red).

Figure 20 shows the mutational profile from sites during week 10, 11, 12 and 13 2023. A combination of spike mutations (V83A, Q183E, Y144del, H146Q, R346T, L368I, F486P and F490S) associated with XBB.1.5 were identified in Hartbeesfontein – Gauteng, Daspoort – Gauteng, Vlakplaats – Gauteng, Rooiwal – Gauteng, Bloemspruit – Freestate, Central eThekwini – KwaZulu Natal and Northern eThekwini – KwaZulu Natal (Figure 20), corroborating with the findings from the Freyja tool, which identified the presence of XBB.1.5 in the same sample (Figure 8). XBB.1.5 was first isolated in South African clinical samples in December, in the Western Cape and continues to emerge in the province and all the other provinces. Wastewater data has shown spike mutations associated with XBB1.5 however the same mutations (V83A, Y144-, H146Q, Q183E, R346T, L368I, F490S) are also associated with BJ.1 (A sub-lineage of BA.2) except for F486P and mutations; T19I, Q23del, L24del, P25del, P26del, I68del, H69del and V70I. Therefore, due to the presence of the other mutations (T19I, Q23del, L24del, P25del, P26del, I68del, H69del, V70I, F486P) in the recent wastewater samples and considering that BJ.1 is a sub-lineage of BA.2 that was only circulating in October, 2022, XBB.1.5 is the lineage that is currently circulating. A combination of mutations (V445P and F486P) in the spike region associated with XBB.1.9 and XBB.1.9.1 were found in Daspoort, Rooiwal, Central eThekwini, Northern eThekwini, Hartbeesfontein and Bloemspruit.

Limitations

The ability to identify variants in wastewater relies on the presence of non-degraded SARS-CoV-2 fragments in wastewater. Our amplicon-based sequencing approach requires binding of primers across the entire SARS-CoV-2 genome. Differential decay of certain portions of the SARS-CoV-2 virus, and disruption of RNA fragments through environmental or chemical pressure leads to imperfect and absent primer binding. In this case, coverage of the genome and the number of reads will be poor or low, and our ability to interpret sequence results and therefore to infer lineages will be impacted.

Conclusion

Qualitative wastewater data from epidemiologic week x, 2023 demonstrate low to moderate levels of SARS-CoV-2 in Gauteng KwaZulu-Natal, Free State and the Eastern Cape. Higher levels have been observed in the Western Cape, corresponding to recent increases in clinical cases. Sequencing data from week 13, 2023 show that recombinant lineages XBB.1.9, XBB.1.9.1 with XBB.1.5 dominance are circulating in March, 2023 in March, 2023 in South Africa. The emergence and significance of XBB.1.9 and XBB.1.9.1 is not yet known, however, lineage XBB.1.9.1 is currently increasing and circulating in Indonesia, South East Asia and Europe. The qualitative and sequencing results must be read along with the SARS-CoV-2 reports generated by the Centre for Respiratory Diseases and Meningitis found at (https://www.nicd.ac.za/wp-content/uploads/2022/03/Update-of-SA-sequencing-data-from-GISAID-18-Mar-2022_2.pdf).

References

Bhoyar, R. C. *et al.* (2021) 'High throughput detection and genetic epidemiology of SARS-CoV-2 using COVIDSeq next-generation sequencing', *PLOS ONE*, 16(2), p. e0247115. Available at: https://doi.org/10.1371/journal.pone.0247115.

Crits-Christoph, A. *et al.* (2021) 'Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants', *MBio*, 12(1), pp. e02703-20.

Gonzalez-Reiche, A. S. *et al.* (2020) 'Introductions and early spread of SARS-CoV-2 in the New York City area', *Science*, 369(6501), pp. 297–301.

Ikner, L. A., Soto-Beltran, M. and Bright, K. R. (2011) 'New method using a positively charged microporous filter and ultrafiltration for concentration of viruses from tap water', *Applied and Environmental Microbiology*, 77(10), pp. 3500–3506.

Khailany, R. A., Safdar, M. and Ozaslan, M. (2020) 'Genomic characterization of a novel SARS-CoV-2', *Gene reports*, 19, p. 100682.

Lara, R. W. I. *et al.* (2020) 'Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing', *medRxiv*.

RC, G. B. R. C. H. (2005) 'Burhans R Elnitski L Shah P Zhang Y Blankenberg D Albert I Taylor J 2005 Galaxy: a platform for interactive large-scale genome analysis', *Genome Research*, 15, pp. 1451–1455.

Acknowledgements

- The contributions of local government and wastewater treatment staff to sample collection and transport is acknowledged and appreciated.
- Students support with sample collections and processing the samples : Mr Thoriso Mooa, Ms Unarine Matodzi, Ms Phiwinhlanhla Nkosi – SAMRC-TB Platform
- The Water Research Commission is thanked for their vision and support.
- The NICD SARS-CoV-2 epidemiology and IT team members are thanked for sharing district and sub-district case burdens in order to generate graphs. These team members include Andronica Moipone Shonhiwa, Genevie Ntshoe, Joy Ebonwu, Lactatia Motsuku, Liliwe Shuping, Mazvita Muchengeti, Jackie Kleynhans, Gillian Hunt, Victor Odhiambo Olago, Husna Ismail, Nevashan Govender, Ann Mathews, Vivien Essel, Veerle Msimang, Tendesayi Kufa-Chakezha, Nkengafac Villyen Motaze, Natalie Mayet, Tebogo Mmaborwa Matjokotja, Mzimasi Neti, Tracy Arendse, Teresa Lamola, Itumeleng Matiea, Darren Muganhiri, Babongile Ndlovu, Khuliso Ravhuhali, Emelda Ramutshila, Salaminah Mhlanga, Akhona Mzoneli, Nimesh Naran, Trisha Whitbread, Mpho Moeti, Chidozie Iwu, Eva Mathatha, Fhatuwani Gavhi, Masingita Makamu, Matimba Makhubele, Simbulele Mdleleni, Tsumbedzo Mukange, Trevor Bell, Lincoln Darwin, Fazil McKenna, Ndivhuwo Munava, Muzammil Raza Bano, Themba Ngobeni.
- The NICD Centre for Respiratory Disease and Meningitis are thanked for their assistance in setting up and troubleshooting PCR testing, and ongoing supportive collaboration.