SARS-CoV-2 Sequencing Update
8 December 2021

Supported by the DSI and the SA MRC
Msoni N. Mlisana K. et al. Lancet Microbe 2020

Prepared by the National Institute for Communicable Diseases (NICD) of the National Health Laboratory (NHLS) on behalf of the Network for Genomics Surveillance in South Africa (NGS-SA)
The genomic data presented here are based on South African SARS-CoV-2 sequence data downloaded from GISAID (www.gisaid.org) on 7 December at 20h00

Data license: https://www.gisaid.org/registration/terms-of-use/


Case data is based on specimen collection date. Cases from https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/weekly-epidemiological-brief/ Test data from https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/weekly-testing-summary/
Number of South African genomes deposited on GISAID, by specimen collection week, 2020 and 2021 (N=24 339*).

Total genomes: 24 339
2020 genomes: 6 381
2021 genomes: 17 958

Sequencing data ending epi week 48 (ending 4 December 2021)
Currently in epi week 49 (ending 11 December 2021)

*This represents the cleaned, de-duplicated dataset of unique sequences. This dataset will be used for all further figures.
GISAID genomes vs total cases, 2020 and 2021 (N=24 339)

All provinces, apart from GP, KZN, NC and WC, have comparable percentage of overall cases and overall sequenced genomes. The majority of November sequencing data is from Gauteng.
Omicron lineage mutation profile

- 45-52 amino acid changes (including deletions) across the whole GENOME
  - 26-32 changes in SPIKE

- Does **not possess** the RdRp G671S change associated with a decrease in Ct value for Delta variants
- Does possess the Δ69-70, which causes the S-Gene Target Failure (SGTF) and was previously seen in the Alpha VOC
Omicron spike mutations compared to other VOC/VOIs

<table>
<thead>
<tr>
<th>Variant</th>
<th>Spike mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omicron</td>
<td>V - - I</td>
</tr>
<tr>
<td>C.1.2</td>
<td></td>
</tr>
<tr>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>Alpha</td>
<td></td>
</tr>
<tr>
<td>Alpha+E484K</td>
<td></td>
</tr>
<tr>
<td>Gamma</td>
<td></td>
</tr>
<tr>
<td>Delta</td>
<td></td>
</tr>
<tr>
<td>Kappa</td>
<td></td>
</tr>
<tr>
<td>Eta</td>
<td>V - -</td>
</tr>
<tr>
<td>Iota</td>
<td>I</td>
</tr>
<tr>
<td>Lambda</td>
<td></td>
</tr>
<tr>
<td>Mu</td>
<td>I</td>
</tr>
</tbody>
</table>

- Multiple changes within the two immunogenic regions in S1 (NTD and RBD)
  - including a three amino acid insertion

- Accumulation of mutations surrounding the furin cleavage site
  - Including combination of N679K and P681H

- Effect of most spike S2 subunit changes have not been defined, but may be linked to immune escape

Mutation frequency key
- ≥ 50% sequences
- 5 - 50% sequences
- <5% sequences

Mutation impact key
- Unknown or unconfirmed impact
- Known immune escape
- Enhanced infectivity
Mutational profile of Omicron sequences

Frequency of Spike SNVs for Omicron (n = 1058)

Frequency of whole genome SNVs for Omicron (n = 1058)

Mutational profile of Omicron is largely shared amongst all sequences
Omicron has been detected in 42 countries across the globe (detections based on GISAID).
Delta dominated South Africa’s third wave with >80% frequency in October, with C.1.2 detection remaining <4%. Omicron appears to dominate November sequencing data but sequencing is ongoing to determine its prevalence.
The Delta variant dominated in September and October in South Africa. Omicron was first detected in South Africa in November, comprising 70% (250/358) of sequences.
C.1.2 has been detected at ≤ 4% of sequences monthly.  
Beta prevalence increased slightly in October and has also been detected in November.  
Omicron first detected in South Africa on November 8th, accounting for 70% (n=250/358) genomes from this month.
Eastern Cape Province, 2021, n = 1673
Free State Province, 2021, n = 1013

FS (N=1013)

Number of genomes

Epiweek

Percent of genomes

Epiweek

Cases
Omicron (21X)
C.1.2 (20D)
Delta (21A)
Delta (21I)
Delta (21J)
Beta (20H. V1)
Alpha (20K. V1)
Kappa (21B)
Eta (21D)
20A
20B
20C
20D
unassigned
198
Gauteng Province, 2021, n = 4797
KwaZulu-Natal Province, 2021, n = 2182
Limpopo Province, 2021, n = 1133
Mpumalanga Province, 2021, n = 1064
Northern Cape Province, 2021, n = 1202
Western Cape Province, 2021, n = 3659
Summary

• New B.1.1.529 (21K) lineage has been designated Variant of Concern Omicron
  • Current earliest detection in South Africa: 8 November, Gauteng
  • Omicron dominates November sequencing data at 70% of genomes (n=250/358). Sequencing is ongoing to determine prevalence of Omicron in other provinces.
  • Omicron has now been detected in 42 countries worldwide

• Delta variant dominated in all provinces until end October
  • The Delta sub-lineages varied by province

• C.1.2 lineage detected in all provinces of South Africa with prevalence of <4% of genomes per month
Additional support and collaborators

**NHLS**  
Koeleka Mlisana  
Zinhle Makatini  
Eugene Elliot  
Florette K. Teurnicht  
Kathleen Subramoney  
Oluwakemi Laguda-Akingba  
Shareef Abrahams  
Greta Hoyland  
Gloria Selabe  
Elias Bereda  
Jeannette Wadula

**Hyrax Biosciences**  
Simon Travers

**Cape Town HVTN Laboratory**  
Erica Anderson-Nissen  
Anneta Naidoo

**Ndlovu Research**  
Hugo Tempelman  
CJ Umunnakwe

**Lancet**  
Allison J. Glass  
Raquel Viana

**Ampath**  
Terry Marshall  
Cindy van Deventer  
Eddie Silberbauer

**Pathcare Vermaak**  
Andries Dreyer  
Howard Newman  
Riaan Writes  
Marianne Wolfaardt  
Warren Lowman

**Bridge-the-Gap**  
Raymond Rott

**Cytespace Africa Laboratories**  
Christa Viljoen

**CAPRISA**  
Salim Abdool Karim  
Nigel Garret

**UKZN - Big Data**  
Francesco Pettruccione  
Ilya Sinayskiy

**University of Oxford**  
José Lourenço

**FioCruz, Brazil**  
Vagner Fonseca  
Marta Giovanetti  
Luiz Carlos Junior Alcantara

**Netcare:**  
Richard Friedland  
Craig Murphy  
Caroline Maslo  
Liza Sitharam

**DSI**  
Glaudina Loots

**SA MRC**  
Glenda Gray

**Africa CDC**  
John Nkengasong  
Sofonias Tessema
National Institute for Communicable Diseases

Centre for Respiratory Diseases & Meningitis
Anne von Gottberg
Thabo Mohale
Daniel Amoako
Josie Everatt
Boitshoko Mahlangu
Noxolo Ntuli
Anele Mnguni
Amelia Buys
Cardia Fourie
Noluthando Duma
Linda de Gouveia
Jackie Kleynhans
Nicole Wolter
Sibongile Walaza
Mignon du Plessis
Stefano Tempia
Mvuyo Makhasi
Cheryl Cohen

Centre for HIV and STIs
Jinal Bhiman
Cathrine Scheepers
Constantinos Kurt Wibmer
Thandeka Moyo
Tandile Hermanus
Frances Ayres
Zanele Molaudzi
Bronwen Lamson
Tandile Hermanus
Mashudu Madzivhandila
Prudence Kgagudi
Brent Oosthuysen
Penny Moore
Lynn Morris

Sequencing Core Facility
Zamantungwa Khumalo
Annie Chan
Morne du Plessis
Stanford Kwenda
Phillip Senzo Mtshali
Mushal Allam
Florah Mnyameni
Arshad Ismail

NICD Groups
NICD COVID-19 response team
NICD SARS-CoV-2 Sequencing Group

Zoonotic arbo and respiratory virus program
Centre for Viral Zoonoses
Department Medical Virology/ NHLS
Tshwane Academic division
University of Pretoria

ZARV research program/UP
Marietjie Venter (Head: ZARV)
Adriano Mendes (Postdoc)
Amy Strydom (Postdoc)
Michaela Davis (MSc, intern medical scientist)

NHLS Tshwane
Prof Simnikiwe Mayaphi (HOD)

Funders:
GIZ/BMBF: African Network for Improved diagnostics and epidemiology of common and emerging infectious agents (ANDEMIA)
G7 Global Health fund, Robert Koch Institute, Dr Fabian Leendertz
Multiple labs from NGS-SA and collaborating private laboratories are contributing to the sequencing effort.
## Variants of Concern (VOC)

<table>
<thead>
<tr>
<th>WHO label</th>
<th>Pango lineages*</th>
<th>GISAID clade</th>
<th>Nextstrain clade</th>
<th>Additional amino acid changes monitored*</th>
<th>Earliest documented samples</th>
<th>Date of designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>B.1.1.7#</td>
<td>GRY</td>
<td>20I (V1)</td>
<td>+S:484K +S:452R</td>
<td>United Kingdom, Sep-2020</td>
<td>18-Dec-2020</td>
</tr>
<tr>
<td>Gamma</td>
<td>P.1</td>
<td>GR/501Y.V3</td>
<td>20J (V3)</td>
<td>+S:681H</td>
<td>Brazil, Nov-2020</td>
<td>11-Jan-2021</td>
</tr>
</tbody>
</table>


*Notable spike (S) amino acid changes under monitoring, which are currently reported in a minority of sequenced samples

*Includes all descendant lineages.

#Includes all Q.* lineages in the PANGO nomenclature system.

§Includes all AY.* lineages in the PANGO nomenclature system.
## Currently designated Variants of Interest (VOI)

<table>
<thead>
<tr>
<th>WHO label</th>
<th>Pango* lineages</th>
<th>GISAID clade</th>
<th>Nextstrain clade</th>
<th>Earliest documented samples</th>
<th>Date of designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lambda</td>
<td>C.37</td>
<td>GR/452Q.V1</td>
<td>21G</td>
<td>Peru, Dec-2020</td>
<td>14-Jun-2021</td>
</tr>
<tr>
<td>Mu</td>
<td>B.1.631</td>
<td>GH</td>
<td>21H</td>
<td>Colombia, Jan-2021</td>
<td>30-Aug-2021</td>
</tr>
</tbody>
</table>


*Includes all descendant lineages.
Submission of routine specimens for sequencing

- representative of multiple geographic regions (provinces/districts/health facilities) from individuals of
  - all ages
  - over as many time periods during the SARS-CoV-2 epidemic in South Africa

- requested that testing laboratories in both the private and public sectors, submit respiratory samples to their closest NGS-SA sequencing laboratory on a routine basis (ideally every week) as follows, depending on the capacity of the testing laboratory:
  - All positives samples should be sent every week (NGS-SA laboratory will perform random sampling as described below) OR
  - A weekly selection of approximately 10%-20% of randomly selected positive samples should be sent every week. Number of selected samples will depend on the size of laboratory and how many other laboratories are drained by the submitting laboratory.
Submission of special interest specimens for sequencing

In addition to routine samples mentioned above, please send specimens separately to above and clearly marked if:

• Suspected vaccine breakthrough (≥14 days after vaccine), especially if hospitalised and clinically severe
• Suspected re-infection (≥90 days after previous episode), especially if hospitalised and clinically severe
• Prolonged shedding with high SARS-CoV-2 viral loads (i.e. Ct values less than 30 for more than 1 month post-primary diagnosis) in immunocompromised individuals
• Possible animal-to-human transmission
• Suspected cases of importation from another country, especially countries known to harbour SARS-CoV-2 variants of concern or countries with little available information
• Clusters of “unusual” cases (e.g., in terms of disease presentation, patient groups affected, etc.)