

SARS-CoV-2 Africa dashboard for real-time COVID-19 information

Joicymara S. Xavier, Monika Moir, Houriiyah Tegally, Nikita Sitharam, Wasim Abdool Karim, James E. San, Joana Linhares, Eduan Wilkinson, David B. Ascher, Cheryl Baxter, Douglas E. V. Pires & Tulio de Oliveira



The SARS-CoV-2 Africa dashboard is an interactive tool that enables visualization of SARS-CoV-2 genomic information in African countries. The customizable app allows users to visualize the number of sequences deposited in each country, and the variants circulating over time. Our dashboard enables near real-time exploration of public data that can inform policymakers, healthcare professionals and the public about the ongoing pandemic.

COVID-19 is the clinical manifestation of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection¹. The COVID-19 pandemic has been ongoing for more than two years, with the first case of COVID-19 in Africa reported in Egypt in mid-February 2020 (ref. ²). The first SARS-CoV-2 genome sequenced in Africa was reported in March 2020 (ref. ³). Genomic sequencing and surveillance have played a crucial role in monitoring and mitigating the COVID-19 pandemic. There have been approximately 12 million cases and more than 256,000 deaths reported to date in Africa⁴, and African countries have contributed substantial amounts of genomic sequencing data to global agencies. For example, two of the five variants of concern (Beta and Omicron) were first identified in Africa through genomic surveillance systems and real-time sequencing and data release^{5,6}.

During the first year of the pandemic, SARS-CoV-2 genomes from Africa were mainly produced for a small number of countries with genomes available from 38 of the 54 African countries⁷. Subsequently, the Africa Centres for Disease Control and Prevention (Africa CDC) and the World Health Organization Regional Office for Africa (WHO AFRO) invested in capacity building and provided resources to equip more African countries to produce genomes locally⁸. For example, the African Union Commission and Africa CDC launched the Africa Pathogen Genomics Initiative (Africa PGI) with an initial investment of US\$100 million. Currently, more than 100,000 genomes, originating from 51 African countries and 4 independent overseas territories, are publicly available from Global Initiative on Sharing Avian Influenza Data (GISAID)⁹.

Dashboards for live COVID-19 information

Online dashboards presenting global and regional COVID-19 data, including case numbers, reported deaths and vaccination rates, have proliferated since the onset of the pandemic^{10–12}. These dashboards have a vital role in guiding the public health response and decision-making by policymakers, public health officials and scientists¹³. Data visualization

in dashboards also keeps the public abreast of the state of the pandemic. Examples of genomic dashboards include the Wellcome Sanger Institute's COVID-19 Genomic Surveillance dashboard (<https://go.nature.com/3U9wS8R>) and the COVID-19 Genomics UK Consortium dashboard (<https://go.nature.com/3Fw32r2>). These dashboards include the number of genomes sequenced and the proportion of variants identified in the sequenced genomes, as well as information on the mutations in the lineages of interest. Although these dashboards display important genomic information about England, there was initially no genomics dashboard for the African continent. We therefore set out to devise a dashboard that provides real-time analytical tools for visualization of a genomics-oriented understanding of the state of the pandemic on the African continent.

Data inputs for the SARS-CoV-2 Africa dashboard

The SARS-CoV-2 Africa dashboard is an open-source web-based graphical user interface for presentation of the data produced by genomic surveillance of COVID-19 on the continent, and for provision of details of variants that are currently circulating. The dashboard is supported by the main commercially available web browsers, including Google Chrome, Mozilla Firefox, Microsoft Edge and Safari. The dashboard collates all sequencing data available in GISAID, with metadata linking data to a specific country in Africa, and uses these metadata to display temporal and spatial trends in SARS-CoV-2 evolution in Africa. Genomics data are incorporated into the dashboard using an application programming interface (API) via an agreement with GISAID. The web application processes it, and includes a data quality assessment that can eliminate poor quality registers – for example, sequences assigned to a variant that was submitted before the variant was identified (Fig. 1a).

Data processing in the SARS-CoV-2 Africa dashboard

SARS-CoV-2 genomes are accessioned on GISAID with contextual metadata (such as patient details, collection and sampling strategies, and sequencing and assembly methods) that are subjected to curation by GISAID before release. GISAID data can be freely accessed and downloaded by users after registration. The data acquisition and processing pipelines use Python 3.6 and the web interface is implemented using Streamlit (<https://go.nature.com/3DqDE3o>), with charts created using Plotly¹⁴. The code can be locally installed for customization in a Conda environment¹⁵. Code and dependencies can be installed by cloning the Github repository, available at: <https://go.nature.com/3WjtMRw>.

Performance of the SARS-CoV-2 Africa dashboard

To evaluate dashboard performance, we implemented an experiment using ApacheBench version 2.3 (<https://go.nature.com/3WjtZEi>) and varying for different levels of concurrency (10, 100, 500 and 1,000 simultaneous access). For each level of concurrency, we performed

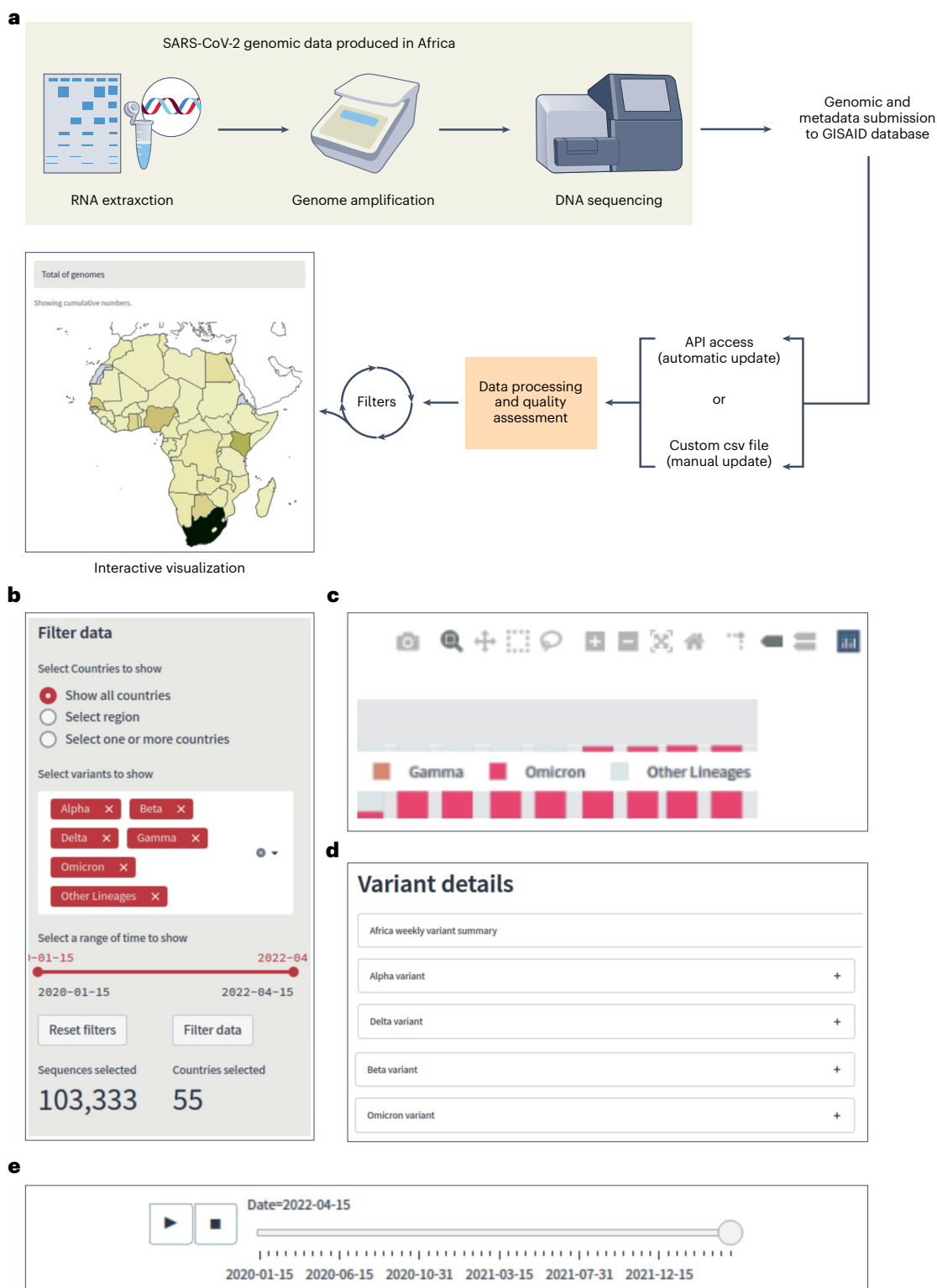


Fig. 1 | Incorporation of SARS-CoV-2 data into the SARS-CoV-2 Africa dashboard. **a**, An overview of the main features of the interface. **b**, General filters that allow users to select the data of interest. **c**, Figure controls that allow users to enable and disable legend elements and labels, select a part of the figure, zoom

in and out, and download the plot. **d**, A tabulated description and mutation map that is provided for variants of interest and/or importance at the time of access. **e**, A timeline player that displays the mapped progression of the pandemic over time, based on the filter selection (**b**).

SARS-COV-2 AFRICA DASHBOARD

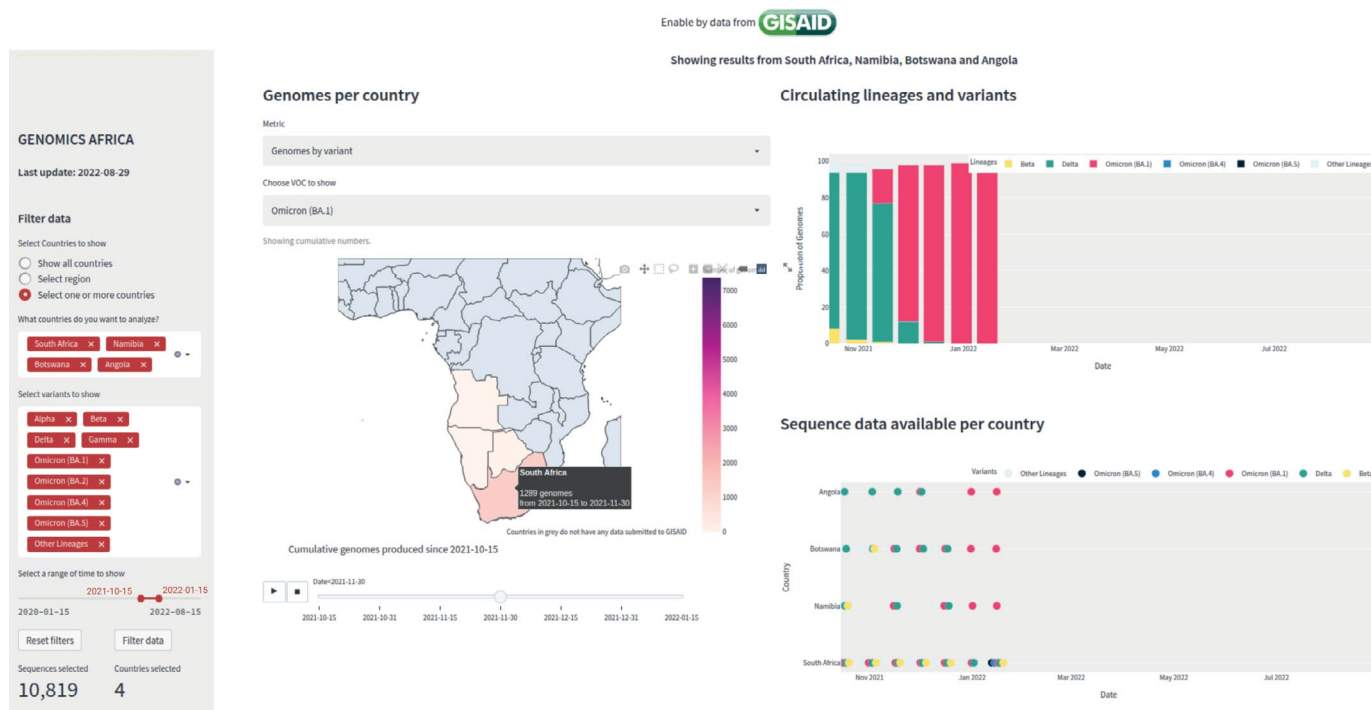


Fig. 2 | A case study of Omicron spread. The case study scenario screenshot shows how to investigate Omicron spread in Namibia and neighbouring countries using the filters and charts provided by the SARS-CoV-2 Africa dashboard interface. Logo courtesy of the GISAID Initiative.

5,000 requisitions, which showed that the dashboard performed well for simultaneous access. For example, in the last level (1,000 requests), only 0.08% of the total requisitions were not answered. All requests for 0–1,000 were completed within an average 1.862 to 25.841 ms (Supplementary Table 1). The dashboard provides interactive visualizations of the temporal and spatial distributions of SARS-CoV-2 variants and their prevalence across different African regions and countries. Several filters are provided to customize the visualizations according to user needs. The main features of the interface are four modules (Fig. 1b–e): general filters allow users to select the data of interest, figure controls allow users to customize the display and snapshot a desired plot, a tabulated weekly summary of variant details is provided, and drop-down mutation maps for variants that are of interest can be used. We also included a timeline player that displays the progression of the pandemic over time, based on user-defined filter selections.

Case study for Omicron spread

A hypothetical example of the application of this dashboard is shown in Fig. 2. The scenario is that the Minister of Health in Namibia wants to understand the spread of the Omicron lineage in neighbouring countries after reports of Omicron in South Africa and Botswana, to better understand how Namibia may be affected. The users would use filters to display the number of Omicron lineage genomes in each neighbouring country. In Fig. 2, on the left-hand side panel of the dashboard, there are filters that allow data for specific countries, specific regions or all countries to be shown. In this example, Namibia and its neighbouring countries (South Africa, Botswana and Angola) have been selected. On the interactive map titled ‘Genomes per country’, the metric ‘Genomes

by variant’ has been selected. In this case, the Omicron variant was selected. As seen in Fig. 2, if the cursor is hovering over a country on the map, the name of the country, the number of genomes produced by that country and the date are displayed.

When studying the figures on the dashboard with these filters applied, one can see that the proportion of genomes deposited in GISAID at the end of October 2021 is dominated by the Delta lineage, with few remnant Beta genomes. Within the first two weeks of November 2021, the Omicron BA.1 lineage rapidly increased to comprise 19% of all genomes. Watching the sliding scale animation for Omicron lineages on the map displays the early detection of the lineage in South Africa, with swift progression from a low (light pink) to high (dark purple) number of cumulative genomes. From these visualizations, the minister would be aware of the rapid spread of Omicron and its growth advantage over Delta, and would be able to see that, at that time, Omicron had the potential to be the dominant variant in southern Africa. The minister would be empowered with the information needed to enable consultation with local researchers, public health officials and clinicians for the provision of local and regional public health responses to mitigate the effects of Omicron on the population.

Outlook


Numerous dashboards for global and regional COVID-19 data, such as case numbers, reported deaths and vaccination rates, have proliferated since the onset of the COVID-19 pandemic. These dashboards have been vital in guiding the public health response and decision-making by policymakers, public health officials and scientists¹³. Data visualizations produced by these dashboards have also been useful for keeping

the public informed. Genomic surveillance of SARS-CoV-2 has been crucial in monitoring the progression of the pandemic, particularly in the low-vaccination landscape of Africa, where globally important variants have emerged and are likely to continue to appear.

Africa has generated a wealth of genomic surveillance data, with more than 129,000 SARS-CoV-2 genomes currently available on GISAID. The SARS-CoV-2 Africa dashboard is the first detailed online, real-time and interactive tool produced for the Global South. It provides simple and clear graphics that are easy to interpret and equips developers to analyse and visualize the data themselves, by allowing manual input of data via custom csv files, formatted as per the provided template (Supplementary Information). Our dashboard makes often intimidating and complex genomic data accessible to all users, and can be used to inform policy and guide the public health response in Africa and for Africa. All datasets used in our dashboard are in publicly accessible repositories. Genomic data are available from the GISAID database (<https://www.gisaid.org/>). The SARS-CoV-2 Africa dashboard is freely available at <https://climade.health/dashboard/covid-africa/>. Source code is available at <https://github.com/CERI-KRISP/SARS-Cov-2-Africa-dashboard>. Supplementary methods are available at <https://doi.org/10.25413/sun.19722025>.

Joicymara S. Xavier ^{1,2,3} , Monika Moir¹, Houriiyah Tegally^{1,4}, Nikita Sitharam¹, Wasim Abdool Karim¹, James E. San ^{1,4}, Joana Linhares¹, Eduan Wilkinson¹, David B. Ascher ^{5,6,7}, Cheryl Baxter ^{1,8} , Douglas E. V. Pires ^{5,6,9}  & Tulio de Oliveira ^{1,4,8,10} 

¹Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, South Africa. ²Institute of Agricultural Sciences, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Unaí, Brazil. ³Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Brazil. ⁴KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), Nelson R. Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa. ⁵Systems and Computational Biology, Bio 21 Institute, University of Melbourne, Melbourne, Victoria, Australia. ⁶Computational Biology and Clinical Informatics, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia. ⁷School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Queensland, Australia. ⁸Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa. ⁹School of Computing and Information Systems, University of Melbourne, Melbourne, Victoria, Australia. ¹⁰Department of Global Health, University of Washington, Seattle, WA, USA.

 e-mail: joicy.xavier@ufvjm.edu.br; cbaxter@sun.ac.za; douglas.pires@unimelb.edu.au; tulio@sun.ac.za

Published online: 22 December 2022

References

- Li, Q. et al. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
- Medhat, M. A. & El Kassas, M. J. *Glob. Health* **10**, 010368 (2020).
- Oluniyi, P. *First African SARS-CoV-2 Genome Sequence from Nigerian COVID-19 Case* (Virological, 2022).
- Ritchie, H. et al. *Coronavirus Pandemic (COVID-19)* (Our World in Data, October 2022).
- Viana, R. et al. *Nature* **603**, 679–686 (2022).
- Tegally, H. et al. Preprint at *medRxiv* <https://doi.org/10.1101/2021.09.23.21264018> (2021).
- Wilkinson, E. et al. *Science* **374**, 423–431 (2021).
- Tegally, H. et al. *Science* <https://doi.org/10.1126/science.abq5358> (2022).
- Khare, S. et al. *China CDC Wkly* **3**, 1049–1051 (2021).
- Wright, D. W. et al. *Virus Evol.* **8**, veac023 (2022).
- COVID-19 Genomic Surveillance* (Broad Institute, October 2022); <https://go.nature.com/3NsMXVi>
- COVID Data Tracker* (CDC, October 2022); <https://go.nature.com/3jcxex>
- Peebles, L. *Nature* **603**, 564–567 (2022).
- Collaborative Data Science* (Plotly Technologies Inc., 2015).
- Anaconda Distribution* (Anaconda Software Distribution, 2016); <https://www.anaconda.com/>

Acknowledgements

We thank GISAID for providing real-time access to metadata available within the database. SARS-CoV-2 sequencing at CERI is supported, in part, by grants from the South African Medical Research Council (SAMRC), World Health Organization, the Rockefeller Foundation, the Abbott Pandemic Defense Coalition (APDC), the National Institute of Health USA (U01 AI151698) for the United World Antivirus Research Network (UWARN) and the INFORM Africa project through IHVN (U54 TW012041), the South African Department of Science and Innovation (DSI) and the SAMRC under the BRICS JAF no. 2020/049. J.S.X. was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brasil (Capes) — finance code 001. D.E.V.P. received funding from an Oracle for Research Grant. Our research included local African researchers from the Centre for Epidemic Response and Innovation (CERI) and the KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP). Local researchers participated in all steps from study design to implementation and authorship (C.B., E.W., H.T., J.E.S., M.M., N.S., T.d.O. and W.A.K.). Participating authors from CERI and KRISP helped to guide the development of the dashboard so that it is locally relevant and useful in Africa. Roles and responsibilities were agreed among collaborators before the start of the project. Guidance on our dashboard has been disseminated to local African researchers via an instructional webinar, as part of the Africa COVID-19 Genomics Training webinar series, hosted by the African Union and Africa CDC. Our research is not restricted to researchers only, and does not pose any health, safety, security or other risks.

Author contributions

J.S.X., H.T., T.d.O., J.E.S. and E.W. devised the project. J.S.X. developed the SARS-CoV-2 Africa dashboard, with contributions from J.E.S., J.L. and W.A.K. H.T., T.d.O., E.W., W.A.K., D.E.V.P., J.E.S., N.S., C.B. and D.B.A. helped with methodology, tests and improvements of the SARS-CoV-2 Africa dashboard and this manuscript. T.d.O. and C.B. managed the project and funding. J.S.X. and M.M. wrote the manuscript, and all authors read and reviewed it.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41564-022-01276-9>.

Peer review information *Nature Microbiology* thanks Amir Bahmani and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.