

Building genomic sequencing capacity in Africa to respond to the SARS-CoV-2 pandemic

Tulio de Oliveira^{1,2,3,4*}, Eduan Wilkinson¹, Cheryl Baxter^{1,3}, Houriiyah Tegally¹, Jennifer Giandhari², Yeshnee Naidoo¹, Sureshnee Pillay²

Genomics surveillance aims to transform public health interventions by monitoring genetic changes that impact pathogenicity, diagnostics, therapeutics, and vaccines. Monitoring the genetic changes in SARS-CoV-2 has played an important role in shaping the scientific response to the pandemic and allowed the identification of several variants of interest (VOI) and five variants of concern (VOC) to date. Although Africa accounts for only about 2.5% and 4.1% of the world's reported COVID-19 cases and deaths, respectively (1, 2), two of the VOC were identified by scientists from South Africa. Here, we reflect on some of the investments and capacity development initiatives that have resulted in an exponential growth

in genomic sequencing capabilities across the continent over the past 2 years.

Early in the SARS-CoV-2 pandemic, genomic surveillance was available in just a few African countries with only 5,245 SARS-CoV-2 genome sequences being made publicly available in 2020 (3). In 2020 and 2021, significant investments in equipment and training were made to extend the geographic coverage of sequencing within many laboratories in Africa, thus increasing surveillance capacity on the continent. These investments resulted in an exponential increase in the number of SARS-CoV-2 genome sequences produced (Figure 1). Interestingly, it took 375 days to produce the first 10,000 SARS-CoV-2 genomes, 87 days to produce the next 10,000, and just 24 days to produce the most recent 10,000 genomes. To date, almost 100,000 SARS-CoV-2 genome sequences from Africa have been shared, and 54 African countries are now contributing to SARS-CoV-2 genome sequencing.

In 2020, the World Health Organization (WHO) and the Africa Centres for Disease Control and Prevention (Africa CDC)

¹Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, South Africa
²KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), Nelson R. Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa
³Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa
⁴Department of Global Health, University of Washington, Seattle, Washington, USA
 *Corresponding author: tulio@sun.ac.za

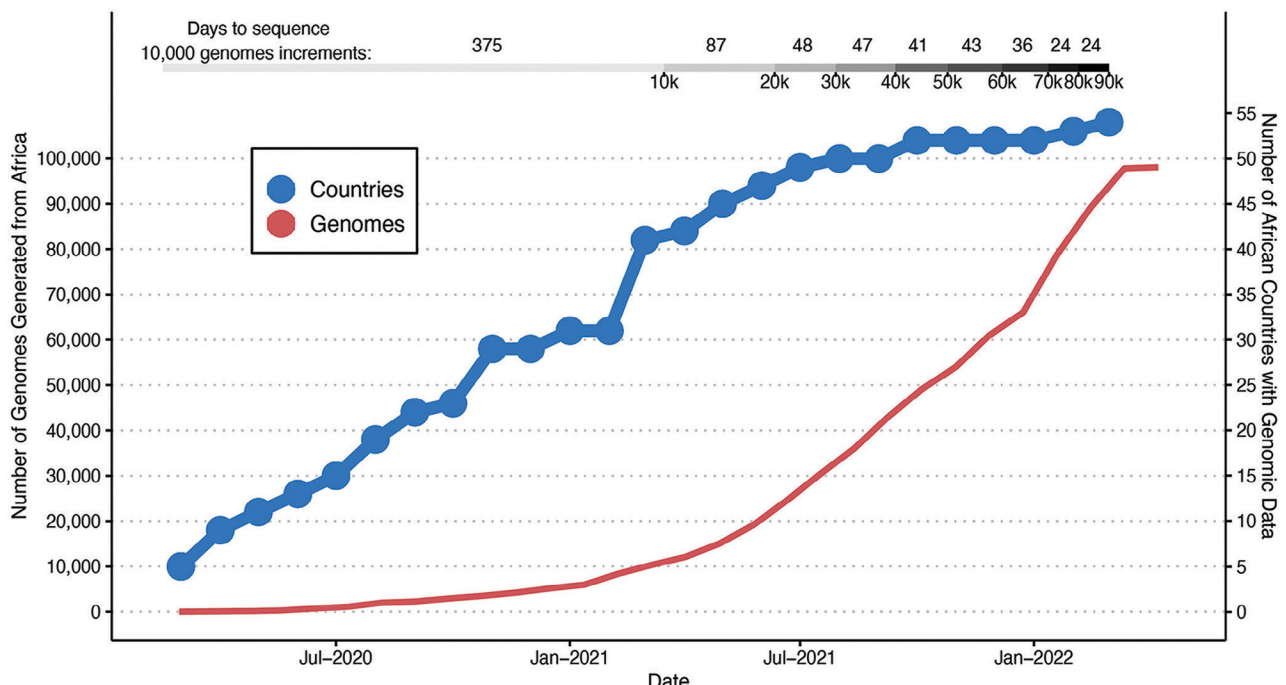


FIGURE 1. Increase in the number of SARS-CoV-2 sequences (solid red line) and African countries (blue circles) contributing genomic sequences to the Global Initiative on Sharing All Influenza Data (GISAID) between January 2020 and March 2022. Note that countries contributing to genomic sequences include 52 African countries and two overseas territories (Reunion and Mayotte).



FIGURE 2. Hands-on genomics sequencing training at the Centre for Epidemic Response and Innovation in Stellenbosch, South Africa.

launched the Pathogen Genomic Initiative (PGI), a network of laboratories established to reinforce SARS-CoV-2 genome sequencing in Africa. Through this initiative and as part of a training program funded by the Rockefeller Foundation and Africa CDC, the Centre for Epidemic Response and Innovation (CERI) and KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP) laboratories in South Africa have provided hands-on genome sequencing and analysis training to 46 fellows from 21 African countries (Figure 2). Importantly, this training has helped build local sequencing and bioinformatics capacity in several of the African countries that were previously sending samples to overseas laboratories for sequencing. In collaboration with the Africa CDC, 18 webinar sessions, which reached 123 African groups, have been hosted to train African researchers in the analysis of their own sequencing data. The first pan-African paper resulting from this collaboration was published in *Science* in 2021 (4), and another publication that highlights the 100,000 genomes generated in Africa is in preparation. An important element of the training is the transparent and open sharing of knowledge. To this end, the protocols established for genomic data

generation have been made publicly available on protocols.io, and all of the data analysis scripts have been shared in GitHub. Future training events are planned through the Rockefeller Foundation and the Abbott Pandemic Defense Coalition, which is dedicated to the early detection of, and rapid response to, future pandemic threats.

The continued generation, analysis, and sharing of virus genomes in real-time from Africa remains important to monitor the expected efficacy and sensitivity of different vaccines and nucleic acid tests across the continent. Africa's increased capacity to not only sequence viruses but also to analyze the data will surely have a multiplier effect and will be beneficial as the continent continues to grapple with ongoing and emergent pathogens in the future.

References

1. E. Dong, H. Du, L. Gardner. *Lancet Infect. Dis.* **20**, 533–534 (2020).
2. Africa CDC, "Africa CDC COVID-19 Dashboard" (2022); available at <https://africacdc.org/covid-19>.
3. Global Initiative on Sharing All Influenza Data (GISAID), "Tracking of Variants" (2022); available at <https://www.gisaid.org/hcov19-variants>.
4. E. Wilkinson *et al.*, *Science* **374**, 423–431 (2021).