A genomics network established to respond rapidly to public health threats in South Africa

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

A genomics network established to respond rapidly to public health threats in South Africa

Published Online
August 18, 2020
https://doi.org/10.1016/S2666-5247(20)30116-6

In response to the COVID-19 pandemic, we created the Network for Genomic Surveillance in South Africa (NGS-SA) in May, 2020 with grants from the South African Medical Research Council and the South African Department of Science and Innovation. Our goal is to sequence the genome of at least 10 000 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) samples to inform the public health response in South Africa. As of July 27, 2020, we have sequenced the genome of 500 isolates, which has provided insights into how and when it was introduced into the country, and information on its early spread through clusters of outbreaks in health-care facilities, workplaces, and other congregate settings.

To respond to the virus, we have taken advantage of the public health research communities that exist in South Africa, and have established a laboratory network that, similar to the COVID-19 Genomics UK Consortium[1] launched in March, 2020, will be guided by the following six key principles: (1) sequencing close to sample collection; (2) being platform agnostic; (3) supporting locally relevant public health priorities, such as by controlling clusters of outbreaks; (4) ensuring rapid and responsible open data sharing; (5) creating a bioinformatics system to process and analyse data locally; and (6) producing timely reports to inform policy makers. In addition, we have embedded a training and capacity-building programme for local scientists and health-care workers.

In order to produce genomic data as close to the point of sample collection as possible, in accordance with the WHO code of conduct,[2] we have linked five large National Health Laboratory Service virology laboratories to a neighbouring academic sequencing centre. This collaboration will allow us to rapidly generate and analyse sequence data to inform regional and national responses near to where the samples are collected and processed.

We have set up protocols for Illumina, Oxford Nanopore Technologies, PacBio, and Ion Torrent sequencers largely based on open methods developed by the ARTICnetwork. We have also modified Illumina protocols to use an alternative library, which has reduced the hands-on sample processing time by 9 h.[3]

As some of the laboratories in the network were prepared when the first cases of COVID-19 were reported in South Africa, we were able to use genomics early in the epidemic to understand the initial spread of the virus, which was characterised by clustered outbreaks. For example, we did a detailed investigation into a nosocomial outbreak in KwaZulu-Natal that began just 4 days after the first reported case in the country. The genomic data supported the hypothesis of a single transmission event that occurred in the emergency department of a hospital, which then led to multiple generations of transmission on several wards within the hospital, and seeded outbreaks in a local nursing home and a renal dialysis unit.[4]

Like the UK consortium,[1] our network in South Africa is committed to the ethical sharing of genomic data, and the corresponding anonymised sample metadata, as soon as they are produced.[2,5,6] Short read data are deposited in the Sequence Read Archive every week. Assembled genomes are rapidly shared via the Global Initiative on Sharing All Influenza Data,[3] which is working in collaboration with the NGS-SA to ensure that data generated in South Africa can also be quality controlled and analysed in-country. We believe that open access to data is crucial to facilitate a rapid response to the SARS-CoV-2 pandemic. We hope that the international scientific community will take this opportunity to partner with us to analyse data in a fair and transparent process that acknowledges and capacitates local researchers and health-care workers.

We have harnessed existing partnerships with the South African National Bioinformatics Institute and the H3Africa Bioinformatics Network to train African bioinformaticians in genome assembly and phylogenetic analysis. We have also produced bioinformatic software applications and protocols to assemble genomic data in research laboratories in South Africa. Genome Detective[7] and Exatype are easy-to-use online software applications. In addition, our bioinformatics team has curated a functional website that provides access to software applications, protocols, data, analytical updates using Nextstrain,[8] and other tools.

For more on the **Network for Genomic Surveillance in South Africa** see https://www.krisp.org.za/ngs-sa/

For more on the **Sequence Read Archive** see https://www.ncbi.nlm.nih.gov/sra

For more on the **Global Initiative on Sharing All Influenza Data** see https://www.gisaid.org

For more on **Exatype** see https://sars-cov-2.exatype.com

For more on the **ARTICnetwork**, see https://artic.network

See Online for appendix

Our work is already influencing policy and practice in South Africa. For example, our genomic analysis of the large nosocomial outbreak has been used to advise health-care facilities across the country on how to strengthen infection prevention and control systems. This analysis was particularly important in South Africa, which had not been affected by previous coronavirus pandemics. We believe that, even though we will not have the resources to sequence the genome of every SARS-CoV-2 sample in the country, targeted and rational use of genome sequencing will be an important resource to support the epidemic response. This will be the first time that genomic epidemiology has been used in real time to inform the public health response to a viral pandemic in South Africa.

We declare no competing interests. The Network for Genomic Surveillance in South Africa is supported by the funding from the South African Medical Research Council and the Department of Science and Innovation.

*Nokukhanya Msomi, Koleka Mlisana, *Tulio de Oliveira, on behalf of the Network for Genomic Surveillance in South Africa writing group†*
deoliveira@ukzn.ac.za

Discipline of Virology (NM), and KwaZulu-Natal Research Innovation and Sequencing Platform (TdO), University of KwaZulu-Natal, Durban, South Africa; National Health Laboratory Services, Johannesburg, South Africa (NM, KM); and Center for the AIDS Program of Research in South Africa (CAPRISA), Durban, South Africa (KM, TdO)

†Members listed in the appendix (p 1).

1 The COVID-19 Genomics UK (COG-UK) consortium. An integrated national scale SARS-CoV-2 genomic surveillance network. *Lancet Microbe* 2020; **1:** e99–100.
2 WHO. WHO's code of conduct for open and timely sharing of pathogen genetic sequence data during outbreaks of infectious disease. 2016. https://www.who.int/blueprint/what/norms-standards/GSDDraftCodeConduct_forpublicconsultation-v1.pdf?ua=1 (accessed May, 10, 2020).
3 Pillay S, Giandhari J, Tegally H et al. Whole genome sequencing of SARS-CoV-2: adapting illumina protocols for quick and accurate outbreak investigation during a pandemic. *Genes* (in press).
4 Nording L. Study tells 'remarkable story' about COVID-19's deadly rampage through a South African hospital. 2020. https://www.sciencemag.org/news/2020/05/study-tells-remarkable-story-about-covid-19-s-deadly-rampage-through-south-african (accessed May 28, 2020).
5 Yozwiak NL, Schaffner SF, Sabeti PC. Make outbreak research open access. *Nature* 2015; **518:** 477–79.
6 Johnson SB, Parker M. The ethics of sequencing infectious disease pathogens for clinical and public health. *Nature Rev Genet* 2019; **20:** 313–15.
7 Cleemput S, Dumon W, Fonseca V, et al. Genome detective coronavirus typing tool for rapid identification and characterization of novel coronavirus genomes. *Bioinformatics* 2020; **36:** 3552–55.
8 Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 2020; **34:** 4121–23.