SARS-CoV-2 sequencing collaboration in west Africa shows best practices

Genomic surveillance can track pathogen transmission during an outbreak, identifying actionable genomic variation on the co-evolution of pathogens with the human genome and socioeconomic demographics.[1] Current data suggest that decentralising sequencing capacity is beneficial, enabling facilities to track local infection dynamics for prompt interventions.[2] Technology has also improved to provide mobile sequencing platforms that are easily deployable even in resource-limited settings. Most importantly, findings from genomic surveillance can guide authorities to establish the best interventions to halt transmission among the populace and reduce the burden on health systems.[3]

The Genomics Unit of the Medical Research Council, The Gambia (MRCG), as part of its programme of building partnerships with research organisations in west Africa, reached out to the Nigerian Institute of Medical Research, Nigeria, to collaborate on whole-genome sequencing of SARS-CoV-2. Since 2008, the Nigerian Institute of Medical Research had acquired capacity for Sanger sequencing, which was used to sequence a fragment and confirm the index case of COVID-19 in Nigeria. However, this method is not cost-effective for genomic surveillance compared with next-generation sequencing methods. Through special government funding for COVID-19 research, the Nigerian Institute of Medical Research procured a next-generation sequencing platform. However, travel restrictions and the logistics limitations associated with the pandemic delayed delivery, installation, training, and next-generation sequencing use. Thus, the collaboration with MRCG was timely and welcome, as the Nigerian Institute of Medical Research sought to build capacity for whole-genome sequencing of SARS-CoV-2. The Centre for Human Virology and Genomics, the Central Research Laboratory at the Nigerian Institute of Medical Research, and the Genomic Unit at MRCG were the key facilities involved.

MRCG proposed whole-genome sequencing using Oxford Nanopore Technology, suitable for rapid deployment in resource-limited settings. MRCG trained Nigerian Institute of Medical Research staff and research interns on the Oxford Nanopore Technology workflow from May 3 to May 11, 2021. 15 participants were trained at the standard four-room sequencing suite at the Centre for Human Virology and Genomics, a HIV drug resistance sequencing facility recognised by WHO. MRCG provided Oxford Nanopore Technology, MK1C equipment, reagents, and consumables. The MK1C can sequence one flow cell per run standalone with 11–96 samples pooled. It does not need an internet connection while running, and the facility only provided power to the equipment. Its minimal size easily fits into the laboratory setup.

Samples were retested to verify their cycle threshold values because they can vary depending on the test kit used.[4] Three sequencing runs of 24 samples each were performed using the Artic nCoV-2019 sequencing protocol v3.[5] We could not manage and analyse the massive fast5 and fastq data produced. The MRCG team performed preliminary data analysis and committed to building local capacity for data analysis, to ensure self-sufficient local teams. We tested 45 samples, and the data suggest a minimum cycle threshold of 30 for samples to sequence successfully. Nextclade assigned clades to 43 sequences, whereas Pangolin assigned lineages to only 29 sequences. 19 sequences had less than 3000 Ns on Nextclade.

Subsequently, a bioinformatics workshop was organised at MRCG, from June 14 to June 25, 2021, with two Nigerian Institute of Medical Research staff participating. The first week dealt with general bioinformatics procedures and software, whereas the second week focused on data analysis specific to SARS-CoV-2 and the Oxford Nanopore Technology reads. It covered quality control, trimming and filtering data, sequence assembly, assembly validation, and variant calling. Participants were introduced to MRCG's high-performance computing cluster, learning to login using a ssh tool, a secure login, and submit data for analysis remotely. Participants received individual logins for future data analysis on the high-performance computing cluster. This capacity building and access to the high-performance computing cluster has enabled local teams to sequence and analyse their data rapidly.

Given the challenge of low internet speed across the subregion, the Nigerian Institute of Medical Research provided a Microsoft computer with graphical processing units processors running Microsoft Windows 10 2.6Ghz 9th generation Intel i7 with 32GB memory RAM and one terabyte solid-state drive hard disk. The MRCG information technology team configured this computer for some level of data analysis locally. It was great learning to configure the Ubuntu operating system, bioinformatics tools, and pipelines on the Windows substation for Linux (WSL2), utilising the graphical processing units and Nvidia CUDA tools. With this computer, we performed base calling on fast5 files skipped in a rush to execute three runs on the minION during the wet-laboratory training.

This engagement typifies excellent South-South collaboration that builds local capacity across the west-African region, and the choice of Nanopore technology was perfect. With its low start-up cost, any facility with the basic laboratory knowledge, equipment, and bioinformatics support can sequence and perform genomic surveillance.

Prompt action to alleviate identified challenges, especially using shared facilities, is crucial for success.

The Nigerian Institute of Medical Research and MRCG intend to expand and affirm their collaboration with a memorandum of understanding to ensure continuity and long-term sustainability. Beyond the pandemic, enduring South-South collaborations should be encouraged.

We declare no competing interests.

Chika Kingsley Onwuamah, Abdoulie Kanteh, Bowofoluwa Sharon Abimbola, Rahaman Ademolu Ahmed, Chika Leona Okoli, Joseph Ojonugwa Shaibu, Ayorinde B James, Olusola Ajibaye, Azuka P Okwuraiwe, Muinah Fowora, Ngozi Otuonye, Archibald Worwui, Bamidele Iwalokun, Dembo Kanteh, Rosemary A Audu, Richard Adebayo Adegbola, Umberto D'Alessandro, Babatunde Lawal Salako, Abdul Karim Sesay

ck.onwuamah@nimr.gov.ng

Centre for Human Virology and Genomics (CKO, BSA, RAAh, CLO, JOS, APO), Biochemistry and Nutrition Department (ABJ, OA), Central Research Laboratory (MF, NO, BI), Microbiology Department (RAAu, RAAd), and Clinical Sciences Department (BLS), Nigerian Institute of Medical Research, Lagos, Nigeria; Genomic Core facility (AK, AKS), HPC facility (AW), Research Support Unit (DK), and Disease Control and Elimination Research Group (UD'A), Medical Research Council Unit, The Gambia at LSHTM, Fajara, Banjul, The Gambia

1 Meredith LW, Hamilton WL, Warne B, et al. Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic surveillance study. *Lancet Infect Dis* 2020; **20:** 1263–71.

2 Nicholls SM, Poplawski R, Bull MJ, et al. CLIMB-COVID: continuous integration supporting decentralised sequencing for SARS-CoV-2 genomic surveillance. *Genome Biol* 2021; **22:** 196.

3 da Silva Filipe A, Shepherd JG, Williams T, et al. Genomic epidemiology reveals multiple introductions of SARS-CoV-2 from mainland Europe into Scotland. *Nat Microbiol* 2021; **6:** 112–22.

4 Onwuamah CK, Okwuraiwe AP, Salu OB, et al. Comparative performance of SARS-CoV-2 real-time PCR diagnostic assays on samples from Lagos, Nigeria. *PLoS One* 2021; **16:** e0246637.

5 Quick J. nCoV-2019 sequencing protocol v3 (LoCost) V.3. 2020. https://protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye (accessed April 30, 2021).