









OPINION ARTICLE

REVISED The rise of pathogen genomics in Africa [version 2; peer review: 2 approved, 2 approved with reservations]

Gerald Mboowa ¹, Francis Kakooza², Moses Egesa³, Stephen Tukwasibwe^{4,5}, Stephen Kanyerezi ^{1,6}, Ivan Sserwadda ^{1,6}, Benson R. Kidenya ⁷, Jupiter Marina Kabahita⁸, Maria Magdalene Namaganda ⁹, Mike Nsubuga¹, Patricia Nabisubi ¹, Alisen Ayitewala⁶, Grace Kebirungi¹, Esther Nakafu¹⁰, Natasha Patience Akwii¹¹

¹The African Center of Excellence in Bioinformatics and Data-Intensive Sciences, Infectious Diseases Institute, College of Health Sciences, Makerere University, P.O. Box 22418, Kampala, Uganda

²Global Health Security, Infectious Diseases Institute, College of Health Sciences, Makerere University, P.O. Box 22418, Kampala, Uganda

³MRC/UVRI & LSHTM Uganda Research Unit, Entebbe, Uganda

⁴School of Medicine, Uganda Christian University, Kampala, Uganda

⁵Infectious Diseases Research Collaboration, Kampala, Uganda

⁶National Health Laboratories and Diagnostics Services, Central Public Health Laboratories, Ministry of Health, P.O Box 7272, Kampala, Uganda

⁷Department of Biochemistry and Molecular Biology, Weill Bugando School of Medicine, Catholic University of Health and Allied Sciences, Mwanza, Tanzania

⁸National Tuberculosis Reference Laboratory/Supranational Reference Laboratory, Plot 106-1062 Butabika Road, Luzira, Uganda

⁹Department of Immunology and Molecular Biology, College of Health Sciences, Makerere University, P.O. Box 77072, Kampala, Uganda

¹⁰College of Veterinary Medicine, Animal Resources and Bio-security (COVAB), Makerere University, Kampala, Uganda

¹¹Department of Plant Sciences, Microbiology and Biotechnology, College of Natural Sciences, Makerere University, Kampala, Uganda







V2 First published: 09 May 2024, 13:468
<https://doi.org/10.12688/f1000research.147114.1>
 Latest published: 04 Sep 2024, 13:468
<https://doi.org/10.12688/f1000research.147114.2>

Abstract

The routine genomic surveillance of pathogens in diverse geographical settings and equitable data sharing are critical to inform effective infection control and therapeutic development. The coronavirus disease 2019 (COVID-19) pandemic highlighted the importance of routine genomic surveillance of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) to detect emerging variants of concern. However, the majority of high-income countries sequenced >0.5% of their COVID-19 cases, unlike low- and middle-income countries. By the end of 2022, many countries around the world had managed to establish capacity for pathogen genomic surveillance. Notably, Beta and Omicron; 2 of the 5 current SARS-CoV-

Open Peer Review

Approval Status    

	1	2	3	4
version 2				
(revision)				
04 Sep 2024	view		view	view
				
version 1				
09 May 2024	view	view		

1. **Erik Bongcam-Rudloff** , Swedish University of Agricultural Sciences, Uppsala, Sweden



2 variants of concern were first discovered in Africa through an aggressive sequencing campaign led by African scientists. To sustain such infrastructure and expertise beyond this pandemic, other endemic pathogens should leverage this investment. Therefore, countries are establishing multi-pathogen genomic surveillance strategies. Here we provide a catalog of the current landscape of sequenced and publicly shared pathogens in different countries in Africa. Drawing upon our collective knowledge and expertise, we review the ever-evolving challenges and propose innovative recommendations.

Keywords

Africa, Genome sequencing, Pathogen Genomic data, Public health, Data sharing



This article is included in the [Pathogens gateway](#).

2. **Ana S Gonzalez-Reiche** , Icahn School of Medicine at Mount Sinai, New York, USA
3. **Gabriel E. Wagner** , Medical University of Graz, Graz, Austria
4. **Onikepe Folarin**, Redeemer's University, Ede, Nigeria

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Gerald Mboowa (gmbwoowa@gmail.com)

Author roles: **Mboowa G:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Kakooza F:** Conceptualization, Funding Acquisition, Investigation, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Egesa M:** Conceptualization, Data Curation, Investigation, Methodology, Resources, Supervision, Writing – Original Draft Preparation; **Tukwasibwe S:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Project Administration, Resources, Writing – Original Draft Preparation; **Kanyerezi S:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Sserwadda I:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Visualization, Writing – Original Draft Preparation; **R. Kidenya B:** Conceptualization, Data Curation, Formal Analysis, Investigation, Resources, Supervision, Validation, Writing – Original Draft Preparation; **Kabahita JM:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Validation, Visualization; **Namaganda MM:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation; **Nsubuga M:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation; **Nabisubi P:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation; **Ayitewala A:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Visualization, Writing – Original Draft Preparation; **Kebirungi G:** Conceptualization, Data Curation, Formal Analysis, Methodology, Project Administration, Visualization, Writing – Original Draft Preparation; **Nakafu E:** Data Curation, Investigation, Methodology, Resources, Visualization, Writing – Review & Editing; **Akwii NP:** Data Curation, Formal Analysis, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: European and Developing Countries Clinical Trials Partnership -TMA2020CDF-3159 Fogarty International Center-U2RTW010672 Public Health Alliance for Genomic Epidemiology - (PHA4GE)INV-038071

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2024 Mboowa G *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Mboowa G, Kakooza F, Egesa M *et al.* **The rise of pathogen genomics in Africa [version 2; peer review: 2 approved, 2 approved with reservations]** F1000Research 2024, 13:468 <https://doi.org/10.12688/f1000research.147114.2>

First published: 09 May 2024, 13:468 <https://doi.org/10.12688/f1000research.147114.1>

REVISED Amendments from Version 1

In the revised manuscript, we have included a section to reflect existing bioinformatics initiatives in Africa. We have also included a discussion on the challenges and problems that can arise when considering ethical and legal issues in genomic collaborations. A new column in Table 1 has been added to include countries and institutions in Africa where specific NGS platforms are located. In Figure 2, we have inserted a plot showing the distribution of SRA data for Africa only from 2013 to date. In the supplementary Table 3 (<https://doi.org/10.6084/m9.figshare.26826301.v1>), we confirm that the table has pathogen data specific to Africa. The revised version also provides estimates that offer insights into the cost implication of building sustainable local bioinformatics capacity in Africa.

Any further responses from the reviewers can be found at the end of the article

Introduction

The global genomic surveillance strategy for pathogens with pandemic and epidemic potential, 2022–2032 was established by the World Health Organisation (WHO) to help countries develop their national genomic surveillance strategy for priority pathogens.¹ The regular collection and sharing of such data are also fundamental for monitoring and effectively responding to outbreaks and for tracking antimicrobial resistance (AMR) to inform decision-making. Furthermore, WHO emphasizes that sharing of such data is critical for preventing, detecting, and timely responding to epidemics and pandemics at all levels, and is in the interest of all Member States.²

Africa remains increasingly prone to infectious disease threats, recording at least 140 disease outbreaks annually^{2,3} besides the high burden of drug resistance that has been estimated to have caused at least 1.05 million deaths in 2019.⁴ This calls for an urgent need to build and sustain capacity for near real-time outbreak detection, characterization, and routine pathogen genomic surveillance. Coronavirus disease (COVID-19) showed that most of the public health institutions in Africa if adequately capacitated and facilitated with high-throughput sequencing platforms, reagents, and data analytics infrastructure, can perform rapid pathogen detection, characterization, surveillance, and equitable data sharing. As of January 4, 2024, over 170,000 Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genomes from 53 out of 55 countries in Africa have been sequenced and shared via GISAID (the Global Initiative on Sharing All Influenza Data).⁵ An important lesson to learn is that we need to facilitate national laboratories and local scientists to undertake routine pathogen sequencing activities within their countries while collaborating with the rest of the globe. The continent still has several pathogens of epidemic and pandemic potential that require near real-time tracking and in response, the Africa Centres for Disease Control and Prevention (Africa CDC) generated a list of priority pathogens for effective emergency preparedness and response.⁶

Pathogen sequencing in Africa

Next-generation sequencing (NGS) is improving how disease outbreaks are accurately detected and investigated at an unparalleled magnitude. Most public health and research laboratories in many countries including Africa have at least one of the common NGS platforms (see Table 1). Different variables affect the choice of NGS platforms to be acquired by an institution such as the cost of the instrument, availability of after-sale service, access and cost of reagents, sequencing run time, throughput, accuracy, strategies for using NGS (targeted/panel, whole-exome, and whole-genome), depth of sequencing coverage, length of sequencing reads, single-end versus paired-end sequencing, genome size of an organism to be sequenced.

COVID-19 flung pathogen genome sequencing to the vanguard of accurate disease outbreak detection, characterization, surveillance of emerging variants, and real-time data sharing. But in the first two years of the pandemic, 78% of high-income settings sequenced >0.5% of their diagnosed COVID-19 cases, compared to 42% of low- and middle-income countries.⁷ Following this inequality, high-income countries were encouraged to support low- and middle-income countries to improve their local sequencing capacity⁷ to ensure an effective global response to the pandemic. Furthermore, a study revealed that sequencing 0.5% of the cases with a turnaround time of less than 21 days could provide a benchmark for SARS-CoV-2 routine genomic surveillance to detect emerging variants.⁷ In Africa, our findings suggest that many countries have built local capacity for pathogen genomics (see Figure 1). This was possible in part because, in the early phase of the pandemic, COVID-19-related travel restrictions forced countries to consider building local capacity for molecular testing and genomic surveillance of SARS-CoV-2 variants other than relying on shipping samples to other countries. By January 4, 2024, the continent had sequenced at least 1.3% of the reported cases (see Table 2) and ranked fourth.

Generally, the reduced costs of genome sequencing have led to the generation of large volumes of genomic sequence data globally as seen from the shared data in the Sequence Read Archive (SRA) for NGS data at the National Center for

Table 1. Key specifications of common NGS throughput platforms in Africa.

Throughput sequencing categories/ Data output range (Gb: gigabyte and Tb: terabyte)	Common equipment/Maximum Read Length (base pairs)/Maximum Raw data Output (Gb: gigabyte and Tb: terabyte)	Key applications/ sequencing strategies	Selected African countries/institutions with specific NGS equipment
Low throughput (1 – 100 Gb)	iSeq 100-2×150 bp - 1.2 Gb MiniSeq - 2×150 bp - 7.5 Gb DNBSEQ-E25-2×150 bp - 7.5 Gb Ion Torrent: Proton/PGM - 400 bp -10 Gb PacBio RS II - 60 kb - 10 Gb MiSeq - 2×300 bp - 15 Gb MinION - 100 kb - 50 Gb DNBSEQ-G99-2×300 bp - 96 Gb	Targeted panel sequencing Pathogen sequencing (microbes and viruses) 16S metagenomic sequencing	<p>MinION (Institut Pasteur Du Maroc, Morocco; National Public Health Reference Laboratory, Somalia; Public Health Laboratory, Chad; Center for Research on Meningitis and Schistosomiasis, CERMES, Niamey, Niger; National Reference Laboratory, Lesotho; Pasteur Institute of Bangui, Central African Republic; University of Science of Technical and Technology De Bamako, Mali; Laboratoire National de Santé Publique de Brazzaville, Republic of Congo; Laboratoire National de Biologie Clinique et de Santé Publique, Central African Republic; Institut Pasteur, Cote d'Ivoire; Jean Piaget Molecular Biology Laboratory, Guinea-Bissau; National Public Health Institute, Cape Verde; Laboratoire d'Analyses Médicales Malagasy, Madagascar, Institut National de Santé Publique, INSP, Burkina Faso; Viral Hemorrhagic Fevers Laboratory of Benin; National Reference Laboratory, São Tomé and Príncipe; Institut National de Santé Publique, INSP, Burkina Faso; Institut Pasteur Dakar, Senegal; Public Health Institute of Malawi, PHIM, Malawi; CPHRL, Sierra Leone; National Institute of Public Health, INSP, Burundi; UVR, Uganda), iSeq 100 (Laboratoire National de Santé Publique de Brazzaville, Republic of Congo; Centre de recherche et de formation en infectiologie de Guinée, CERFIG, Guinée; Laboratoire National de Référence, Togo; INRB, DRC; Institut Pasteur Dakar, Senegal; Public Health Laboratory, Seychelles; UVR, Uganda), MiniSeq (Pasteur Institute of Bangui, Central African Republic; Central Health Laboratory, Mauritius; Rwanda Biomedical Centre, RBC, Rwanda; Public Health Institute of Malawi, PHIM, Malawi; CPHRL, Sierra Leone; National Institute of Public Health, INSP, Burundi), MiSeq (National Reference Laboratory, Lesotho; Jean Piaget Molecular Biology Laboratory, Guinea-Bissau; Instituto Nacional de Saúde Pública, INSP, Angola; NCDC, Nigeria; CPHL & UVR, Uganda; EPHI, Ethiopia; CPHL, Egypt; INRB, DRC; NICD, South Africa; National Public Health Laboratory, Tanzania; ILRI & KEMRI, Kenya; NPHL & National University Research Institute, Khartoum, Sudan; Institut Pasteur d'Algérie, Algeria; Institut Pasteur Du Maroc, Morocco; Instituto Nacional de Saúde, INS, Mozambique; Noguchi Memorial Institute for Medical Research,</p>

Table 1. *Continued*

Throughput sequencing categories/ Data output range (Gb: gigabyte and Tb: terabyte)	Common equipment/Maximum Read Length (base pairs)/Maximum Raw data Output (Gb: gigabyte and Tb: terabyte)	Key applications/ sequencing strategies	Selected African countries/institutions with specific NGS equipment
Medium throughput (100 – 1000 Gb)	<p>NextSeq - 2 × 150 bp - 120 Gb</p> <p>DNBSEQ-G50RS - 2 × 150 bp - 150 Gb</p> <p>GridION - 100 kb - 250 Gb</p> <p>PacBio Sequel II/IIe - 60 kb - 500 Gb</p>	<p>Small whole-genome sequencing</p> <p>Targeted panel sequencing (amplicon and gene panel)</p> <p>Gene expression profiling with mRNA-Seq</p> <p>miRNA and Small RNA analysis</p>	<p>MMIMR, Ghana; Institut Pasteur, Madagascar; Le Laboratoire National de Santé Publique, LNSP, Cameroon; Zambia National Public Health Institute, ZNPHI, Zambia; National Microbiological Reference Laboratory, NMRL, Zimbabwe; Institut Pasteur de Tunis, Tunisia; National Public Health Reference Laboratory, Liberia; NIP & UNAM, Namibia; MRC Unit, The Gambia; Centre Interdisciplinaire de Recherches Médicales de Franceville, CIRMF, Gabon; Botswana Harvard AIDS Institute Partnership, Botswana; Institut National de Santé Publique, INSP, Burkina Faso; National Institute of Public Health, INSP, Burundi; Institut Pasteur Dakar, Senegal)</p>
High throughput (1000 – 10 000 Gb)	<p>DNBSEQ-G400RS - 2 × 300 bp - 1.4 Tb</p> <p>HiSeq - 2 × 150 bp - 1.5 Tb</p> <p>DNBSEQ-T7-2 × 150 bp - 7 Tb</p>	<p>Whole-genome sequencing</p> <p>Whole-exome sequencing</p> <p>Transcriptome sequencing</p> <p>Shotgun metagenomics</p>	<p>GridION (Laboratoire National de Santé Publique de Brazzaville, Republic of Congo; KEMRI-Wellcome Trust Research Programme, Kenya; INRB, DRC; Rwanda Biomedical Centre, RBC, Rwanda; Institut Pasteur Dakar, Senegal), NextSeq (Public Health Institute of Malawi, PHIM, Malawi, National Public Health Laboratory, Eswatini; Institut Pasteur Du Maroc, Morocco; Noguchi Memorial Institute for Medical Research, MMIMR, Ghana; ACEGID, Nigeria; INRB, DRC; National Institute of Public Health, INSP, Burundi; EPHI, Ethiopia; Institut Pasteur Dakar, Senegal), PacBio Sequel II (ACEGID, Nigeria)</p>
Ultra-high throughput (>10 Tb)	<p>PromethION - 270 kb - 13.3 Tb</p> <p>NovaSeq - 2 × 250 bp - 16 Tb</p> <p>DNBSEQ-T20 × 2RS - 2 × 150 bp - 72 Tb</p> <p>DNBSEQ-T10 × 4RS - 2 × 150 bp - 76.8 Tb</p>	<p>Ultra-high-depth whole-genome sequencing of large genomes (Human, plant and animal)</p> <p>Whole-exome sequencing</p> <p>Whole-transcriptome sequencing</p> <p>Methylation sequencing</p> <p>Shotgun metagenomics</p>	<p>NovaSeq X (ACEGID, Nigeria), NovaSeq 6000 (CERI, South Africa; Institut Pasteur Dakar, Senegal)</p>

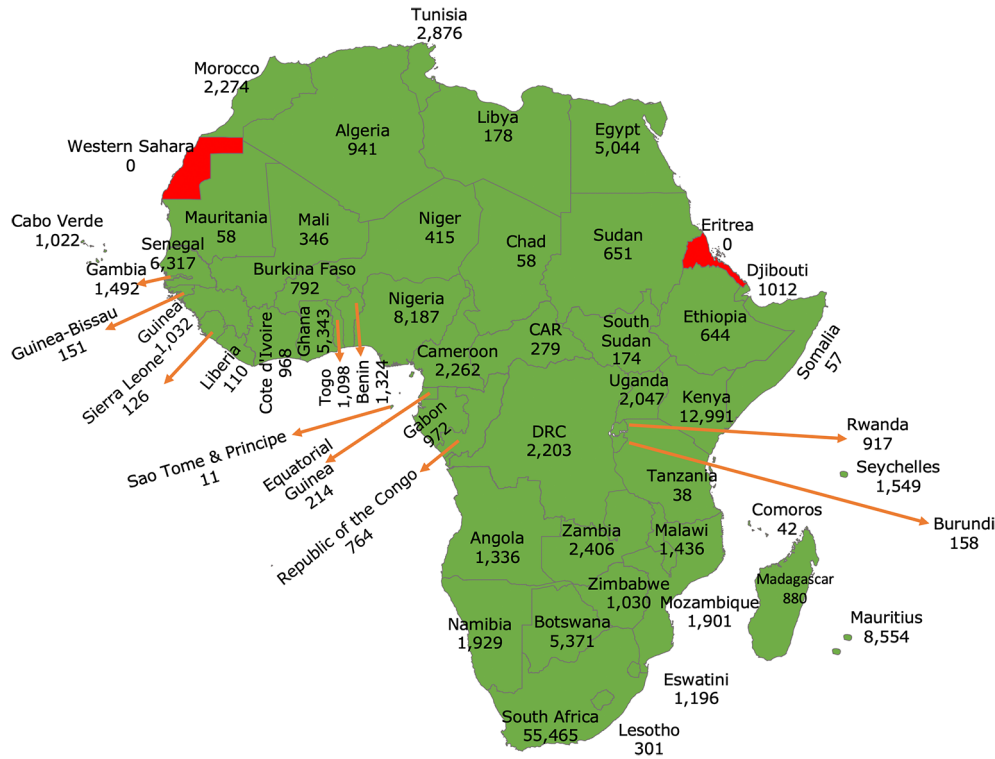


Figure 1. Map of Africa showing the number of SARS-CoV-2 genomes sequenced and shared via GISAID per country. (GISAID data accessed on January 4, 2024). A total of 53/55 African Union Member States sequenced and shared their SARS-CoV-2 genomes since the start of the pandemic. At least 96% sequenced by Hubs based in Africa.⁸

Table 2. Shows proportions of SARS-CoV-2 cases sequenced by each continent. Reported number of COVID-19 cases* and SARS-CoV-2 sequence data shared via the Global Initiative on Sharing All Influenza Data (GISAID)[#] by each continent. Data accessed on January 4, 2024.

Continent	COVID-19 cases	Sequenced SARS-CoV-2 genomes in GISAID	Percentage of COVID cases sequenced & shared
North America	130,243,977	5,781,894	4.40%
Europe	252,605,493	8,027,917	3.20%
Oceania	14,722,723	283,139	1.90%
Africa	12,857,043	171,881	1.30%
Asia	221,152,538	1,701,264	0.80%
South America	69,454,296	398,332	0.60%

*(https://www.worldometers.info/coronavirus/?utm_campaign=homeAdUOA?Si%23countries).

[#](<https://gisaid.org/>).

Biotechnology Information (NCBI) (see Figure 2). The SRA was established as part of the International Nucleotide Sequence Database Collaboration (INSDC) in 2009.⁹

It should be noted that more genomic data is being generated than the capacity to analyze and meaningfully interpret it.^{10,11} This data if well utilized holds clues to improve health. This makes FAIRness (findability, accessibility, interoperability, and reusability) data sharing¹² a critical component in this ecosystem so that scientists in well-resourced settings can collaboratively and equitably derive valuable insights from this data.

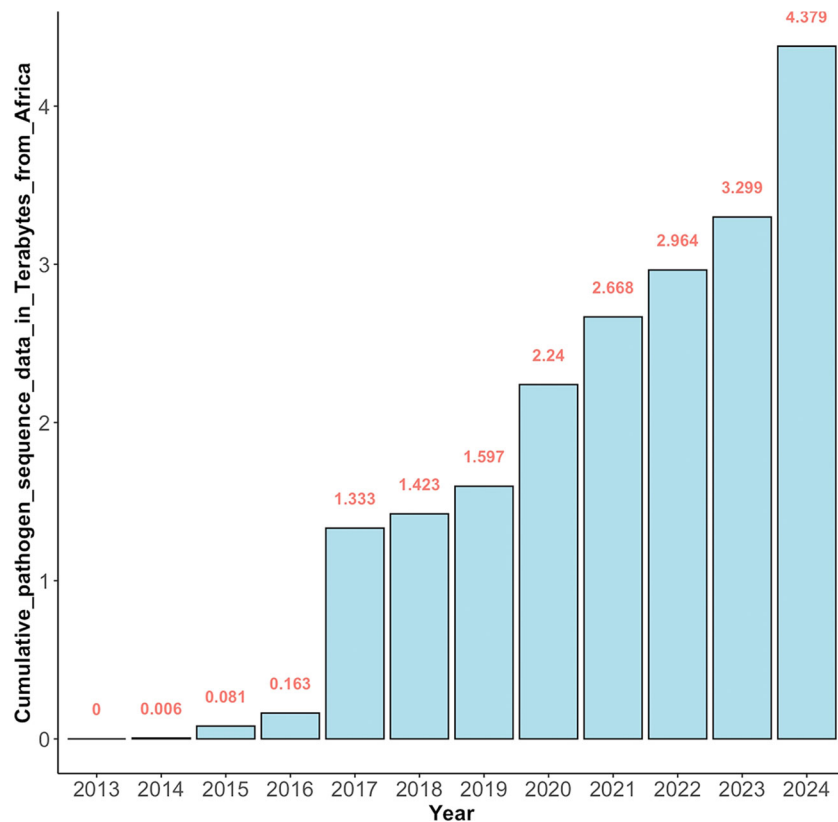


Figure 2. The cumulative amount of pathogen genomic sequence data in Terabytes deposited in SRA since 2013. Over the span of 12 years (2013-2024), at least 4.4 terabytes of pathogen genomic sequence data from Africa have been deposited in the Sequence Read Archive (SRA), making it publicly accessible.

Africa's urgent need to develop a robust network of skilled bioinformatics researchers

Africa, with its rich biodiversity, stands to gain most from advancements in biomedical research and healthcare. However, harnessing this potential requires a critical mass of researchers skilled in genomics and bioinformatics. These fields combine biology, computer science, and information technology to comprehensively analyze and interpret vast biological data. Building a critical mass of this expertise in Africa is imperative for leveraging the continent's biodiversity, addressing health challenges, and promoting sustainable development. Existing initiatives like AI4D Africa (<https://africa.ai4d.ai/>), H3ABioNet (<https://www.h3abionet.org/>), EANBIT (<https://eanbit.icipe.org/>), African BioGenome Project (<https://africanbiogenome.org/>) and several university training programs provide a strong foundation, but there is a need for intensified broad training in Artificial intelligence (AI), High performance computing (HPC), NGS sequencing, and bioinformatics analysis. By adopting collaborative strategies, increasing funding, and leveraging technology, the continent can cultivate a skilled workforce in these areas, driving scientific innovation and improving public health outcomes.

Bioinformatics is pivotal in modern science, underpinning advancements in genomics, and personalized medicine. For Africa, developing bioinformatics capacity is crucial due to several reasons such as; (i) Africa faces a high burden of endemic infectious diseases like malaria, HIV/AIDS, and tuberculosis, alongside raising cases of non-communicable diseases, (ii) bioinformatics enables the analysis of genomic data to understand pathogen evolution, drug resistance mechanisms, and host-pathogen interactions, thereby facilitating the development of effective treatments and vaccines, and (iii) Africa's great genetic diversity offers a unique resource broadly for understanding human evolution and disease susceptibility. Therefore, bioinformatics is essential for analyzing Africa's complex genomic data, uncovering genetic markers, and tailoring healthcare solutions to diverse populations.

We propose the following strategies for enhancing training;

1. Establish collaborative training programs through partnerships between African institutions and international institutions to deliver joint training programs. These programs can offer access to cutting-edge technologies, expert mentorship, and global research networks.

2. Utilize online platforms to deliver bioinformatics courses and workshops. This approach can reach a broader audience and provide flexible learning opportunities.
3. Avail funding and scholarships opportunities for bioinformatics education and research. Scholarships and grants can support students and researchers, enabling them to pursue advanced studies and attend international conferences and workshops.
4. Foster collaborations between academia and industry to provide practical training, innovations and research opportunities. Industry partners can offer internships, project collaborations, and access to proprietary technologies.

Data sharing of pathogen genomic data

COVID-19 propelled two important components in effective infectious disease response; (i) near-real-time pathogen sequencing and (ii) unprecedented collaborative scale in pathogen data sharing with over 16.4 million SARS-CoV-2 genomes already publicly shared via GISAID from all over the world by January 4, 2024.

Since most of the SARS-CoV-2 sequence data was shared via GISAID, we searched the SRA/NCBI database (<https://www.ncbi.nlm.nih.gov/sra/>) up until January 4, 2024, for publicly accessible sequence data for other pathogens from different countries in Africa. The search terms used were, “Pathogen [All Fields] AND Country [All Fields]”. The retrieved data included different pathogens sequenced from each country filtered by microbial taxon, BioProject, genomic library layout, NGS platform, volume of data, and submitting/sequencing institution (see Extended data: Table 3¹³).

In Africa before the COVID-19 pandemic, sequencing was largely a sophisticated endeavor undertaken by individuals and institutions based in high-income settings in collaboration with African institutions (see Extended data: Table 3¹³). The local institutions largely participated in sample collection and processing for shipping to institutions in well-resourced countries. This implied that the objectives of these sequencing activities were also addressing the research aims of institutions in high-income settings. The pandemic brought about a paradigm shift that allowed local public health institutions to mobilize resources to embrace the genomic revolution in public health. This has been largely achieved through the Africa Pathogen Genomics Initiative (Africa PGI),³ which has equipped, trained, and offered technical support for African Union (AU) Member States to build sustainable genomics and bioinformatics capacity. Furthermore, the AU Commission and Africa CDC have called on governments, multi-lateral organizations, philanthropies, the private sector, and civil society organizations to support the full implementation of Africa’s New Public Health Order to drive global health security.¹⁴ In addition to this, it is imperative to promote equitable data sharing for pathogens of epidemic and pandemic potential in Africa while discouraging post-outbreak sanctions on countries that share data.

Major emerging challenges

The price of many of the commonly used sequencing equipment widely distributed in Africa varies significantly. For instance, the MinION Mk1B costs \$1,999 with no annual fee, while the MinION Mk1C (which includes a sequencer, display, and GPU/CPU) is priced at \$4,900 plus a \$300 annual fee. The GridION is priced at \$49,955 with an additional yearly cost of \$12,500. Each MinION flow cell yields between 5-20 Gb per run and costs between \$475 to \$900, depending on the quantity purchased.^{15,16} Illumina platforms have a broader price range, starting from \$19,900 for the iSeq 100, \$49,500 for the MiniSeq, and \$128,000 for the MiSeq. Higher-end models include the NextSeq 1000 at \$210,000 and the NextSeq 2000 at \$335,000.

A recent survey indicates that setting up a bioinformatics computing workstation in an African laboratory, particularly one involved in routine pathogen sequencing and bioinformatics analyses, is estimated to cost a minimum of \$7,000. This estimate includes the cost of a desktop computer with high-end specifications such as an Intel Core i9 processor with 16 cores at 5.0 GHz, 16 GB DDR4 RAM, a 12 GB GPU, dual 4 TB M.2 SSDs, and an additional 8 TB internal SSD for storage. Additional expenses, including electricity and internet, must also be considered. The average annual salary for a skilled bioinformatician in a public health laboratory in Africa is approximately \$9,600.

The cost of bioinformatics analysis for a pathogen genome on a cloud platform can vary depending on several factors, including the type of analysis (e.g., pathogen identification, virulence, and antimicrobial resistance profiling), the sequencing strategy used (such as metagenomics or amplicon-based sequencing), and the size of the pathogen genome. Excluding costs related to internet connectivity and data storage, the cost of analyzing a pathogen genome can range from \$0.5 to \$10 per sample, depending on the bioinformatics cloud platform service provider. This estimate typically covers essential tasks like sequence quality assessment, genome assembly, species or lineage assignment, genome annotation,

and variant calling. The runtime and cost can vary depending on the software used, as different tools utilize different algorithms. More complex analyses, such as comparative genomics and the determination of antimicrobial resistance and virulence factors, may increase costs due to the need for additional public pathogen sequences and reference databases. Additionally, the choice of bioinformatics tools is influenced by factors such as the sequencing approach used, including metagenomic next-generation sequencing, whole-genome sequencing, targeted sequencing (amplicon sequencing), multiplex PCR-based NGS, and whether short or long-read sequencing technologies are employed.

The ever-growing volume of genomic sequence data coming out of laboratories in Africa has amplified interest in genomics and bioinformatics applications in public health.¹⁷ However, majority of the workforce has limited experience in these fields largely due to the limited number of training institutions that can offer the training. It should be noted that meaningful interpretation of genomic data generated from disease outbreaks requires adequate training and collaborative efforts (within and outside the continent). Sequencing facilities have varying levels of expertise in sequencing different organisms such as viruses, bacteria, fungi, and parasites as well as different sequencing technologies (see [Table 1](#)).

There is an urgent need for public health institutions in Africa to recruit and retain competent personnel dedicated to performing bioinformatics analyses that can support evidence-informed public health decision-making. In some cases, it has been complicated to establish such job positions within institutional structures. As such resorting to training the available personnel in this field is the most reasonable option however, this training requires long periods of learning, mentorship, and hands-on practice.¹⁸

Limited local data analytics infrastructure is another critical emerging challenge. With many institutions having access to local NGS sequencing platforms that generate large volumes of raw sequence data that require memory-intensive analytic compute and storage, there is an ever-growing need to provide access to local high-performance computing (HPC) facilities. Setting up HPC facilities requires a lot of money and creates new challenges such as increasing electricity consumption and a need to recruit a competent HPC systems administrator. Furthermore, utilization of cloud-based resources for bioinformatics analyses also requires access to stable fast internet that may be unavailable in some countries.

The need to invest in biobanking and biorepository services in Africa. These facilities archive and distribute well-characterized biospecimens for research to support the development of disease diagnostics and therapeutics.¹⁹ Genomics activities rely on access to samples in these facilities. Many public health laboratories faced the challenge of storing thousands of samples during 2014–2016 Ebola outbreak in West Africa and the COVID-19 outbreak and realized a need to expand their biobanking capacities.

Challenges of timely access to NGS reagents and supply chain. Travel restrictions during COVID-19 highlighted the need to improve the supply chain of NGS reagents in countries that lacked or had limited access to channel partners. Routine genomic surveillance requires reliable access to affordable NGS reagents and after-sales service. It is estimated that the cost per pathogen sequence ranges from US \$20–\$200, with poor countries paying up to 10-fold more than high-income countries.²⁰

Pathogen genomics in public health in Africa requires continuous engagement among different stakeholders including NGS platform manufacturers/distributors, program/project funders, sequencing facilities, research/academic institutions, the national ministries of health, and policymakers. Each of these has a unique contribution, therefore planning and supporting multi-pathogen activities is a multi-sectoral endeavor that requires continued engagement to ensure an impactful genomic ecosystem. For example, the prohibitive cost of using onboard NGS instrument data analysis software licenses requires engagement with platform manufacturers. Cloud-based computing is an enthralling emerging infrastructure that offers an alternative infrastructure for solving traditional bioinformatics challenges in NGS data analytics and visualization. Cloud computing service models are classified into three types: Platform as a Service (PaaS), Software as a Service (SaaS), and Infrastructure as a Service (IaaS).²¹ The platform is relatively easier to use than the Linux command-line-based environment however it requires a pay-as-you-go subscription that charges based on usage.

Recent years have revealed an emerging dimension of the impact of climate change on the evolution of pathogens in Africa.^{22,23} For example, climate change and extreme weather conditions have triggered increased cholera outbreaks in Africa.²⁴ Research on modeling climate change and pathogen genomic data has failed to be socially inclusive, geographically balanced, and broad in terms of the disease systems studied, limiting our capacities to better understand the actual contribution of climate change on disease outbreaks.²² Therefore, integrating climate change data in pathogen genomics is important, and leveraging this expertise through collaboration with institutions that have established capacity.

Challenges of integrating ethics and law in international collaborations

International collaboration is essential in genomic activities and global health security, fostering innovation, access to sequencing technologies, sharing knowledge, and addressing complex issues that transcend borders such as during disease outbreaks. However, such collaborations often face significant challenges, particularly when it comes to integrating ethics and law across diverse national legal systems. Ethical guidelines in Africa need adaptation to the evolving data sharing policy landscape, which increasingly emphasizes openness, storage, sharing, and fair secondary use of data. Some countries have guidelines that do not adequately address the ethical challenges these new orientations present, failing to provide accurate guidance to ethics committees and researchers.²⁵

Furthermore, pathogen genomics involves obtaining samples from biobanks. Several countries have established biobanks and biorepositories to support genomic activities in Africa.^{19,25–28} Access to these facilities requires balancing data privacy, trust, and a need to promote scientific progress. Biobanking services involving large prospective cohorts have addressed these issues through extensive communication among collaborators, researchers, community engagement and experts. Funders' main concerns include privacy, biosamples utilization, and information access. Emphasis is often on privacy and funder's control over biosample use. To avoid delays, it's crucial to address ethical and legal issues proactively.²⁹

Integrating ethics and law in international collaborations presents significant challenges due to differences in ethical standards, legal frameworks, cultural values, and enforcement mechanisms. Harmonizing data standards, establishing clear agreements, joint review mechanisms, regulatory capacity building, and fostering cultural competence are essential to navigating these complexities. Legal systems across Africa are diverse, reflecting varied political, economic, cultural, and social structures. This diversity complicates regional and international collaborations, particularly in areas like public health genomics, intellectual property, and data protection. Conflicting laws between African countries require meticulous legal analysis and negotiation. Compliance with multiple legal frameworks can be resource-intensive and complex, with liability in legal disputes further complicated by jurisdictional differences. Regulatory approval for multi-country genomics research in Africa is often cumbersome due to differing requirements and timelines, leading to project delays and increased costs. However, mechanisms established by the Africa Centres for Disease Control and Prevention (Africa CDC) have begun to address these challenges by supporting member states in strengthening their health institutions and navigating the legal complexities of disease threats.

Beyond the post-COVID-19 genomics investment in Africa and way forward

COVID-19 allowed most of the countries in Africa to build local capacity for pathogen genome sequencing. It is important to keep the momentum by leveraging this investment to fight the continent's endemic health challenges including Malaria, Cholera, HIV/AIDS, Tuberculosis, Vaccine-preventable diseases, Viral hemorrhagic fevers (VHFs) and the growing threat of antimicrobial resistance. Therefore, as recommended by the WHO,³⁰ countries are establishing multi-pathogen genomic surveillance strategies that will ensure they can rapidly detect and precisely characterize pathogens, but this also requires stable funding from national governments and other funding agencies such as the Pandemic Fund.

The global economic cost of the COVID-19 pandemic was estimated at USD \$16 trillion.³¹ Much as countries now appreciate the benefits of routine pathogen surveillance, many in Africa require technical and financial support. The devastating human, economic, and social cost of COVID-19 highlighted the urgent need for well-coordinated action to establish stronger health systems and mobilize additional resources for pandemic prevention, preparedness, response, and serve as a platform for advocacy. The WHO coordinated efforts through the newly established Pandemic Fund, which is a collaborative partnership among donor governments, co-investor countries, foundations, civil society organizations, and international agencies providing a dedicated stream of long-term funding (see [Figure 3](#)) for critical pandemic prevention, preparedness, and response (PPR) for future pandemics. The Pandemic Fund is set to finance critical investments to strengthen pandemic prevention, preparedness, and response capacities at national, regional, and global levels, with a focus on low- and middle-income countries.³²

The Pandemic Fund is set to allocate financing where investments are most urgently needed to bolster PPR for future pandemics, addressing key capacity gaps at national, regional, and global levels. In 2023, projects that were financed prioritized strengthening comprehensive outbreak surveillance and early warning, laboratory systems, and human resources/public health workforce capacity. Overall, the Fund's aims include: (i) supporting surveillance systems that enable timely tracking and reporting of outbreaks, (ii) creating faster and more accurate data sharing and (iii) building better laboratories so that they better assist partner countries to rapidly detect and effectively respond to infectious disease outbreaks.

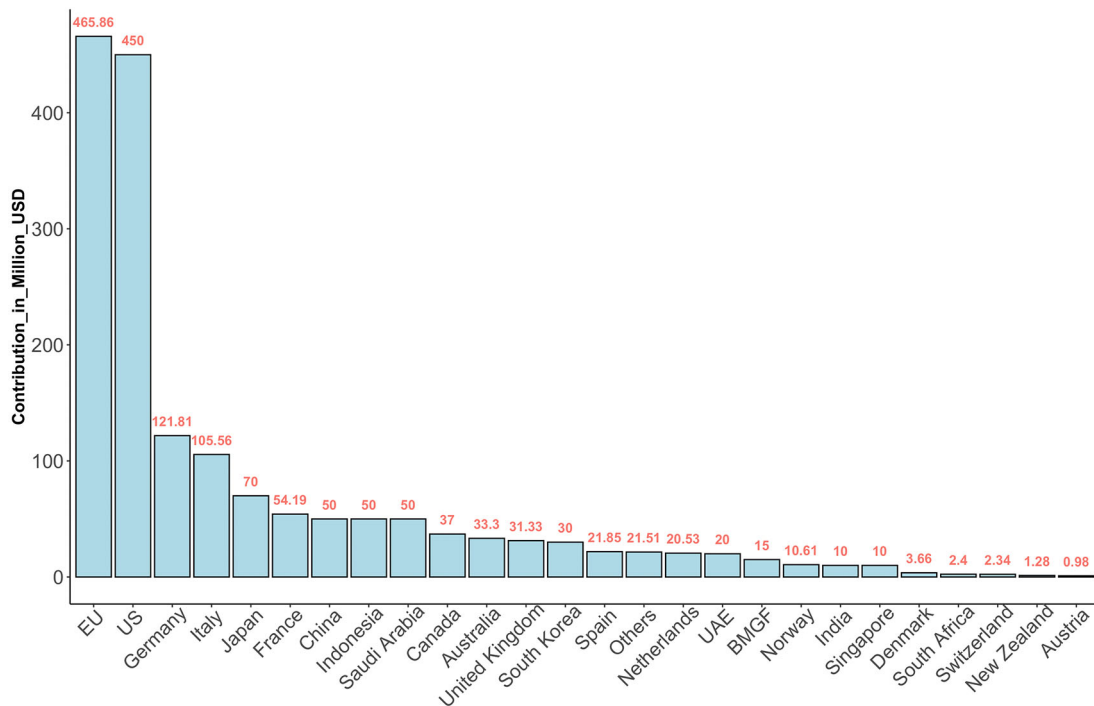


Figure 3. Funds contributed to the Pandemic Fund. The Fund aims to help countries build resilience to future pandemics. The data was accessed on January 4, 2024 (<https://fiftrustee.worldbank.org/en/about/unit/dfi/fiftrustee/fund-detail/pppr>).

Beyond analysis, reporting, and data sharing of pathogen genomic data, the continent also needs to utilize the genomics resource to improve local research and development (R&D) capacity as well as support local therapeutics and affordable diagnostic manufacturing. The Partnerships for African Vaccine Manufacturing (PAVM) has a goal to enable the African vaccine manufacturing industry to develop, produce, and supply over 60 percent of the total vaccine doses required to fight endemic diseases on the continent by 2040.³³ This offers an exceptional opportunity for enhancing local vaccine training. Rwanda is a pioneering country, hosting the first mRNA vaccine manufacturing facility in Africa, and soon, Senegal, South Africa, and Kenya will be the other African countries to have mRNA vaccine manufacturing facilities on the continent.³⁴

Conclusions

Timely and accurate analysis of pathogen genomes is crucial for monitoring the molecular evolution and dissemination of pathogens, improving diagnostic molecular tests and vaccines, and guiding efficient public health interventions. Therefore, there is an urgent need to sustain gains in local pathogen genomics in Africa while strengthening data analytics infrastructure and an equitable data sharing ecosystem.

Data availability

Underlying data

No data are associated with this article.

Extended data

Zenodo: Extended data for ‘The rise of pathogen genomics in Africa’, <https://doi.org/10.6084/m9.figshare.26826301.v1>.¹³

This project contains following extended data:

Table 3. The different pathogens that have been sequenced and publicly shared via SRA:NCBI from different countries in Africa (at least 12 Terabytes of sequence data).xlsx

Data are available under the terms of the [Creative Commons Attribution 4.0 International license \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Acknowledgements

G.M. acknowledges the EDCTP2 career development grant which supports the Pathogen detection in HIV-infected children and adolescents with non-malarial febrile illnesses using the metagenomic next-generation sequencing approach in Uganda (PHICAMS) project which is part of the EDCTP2 programme from the European Union (Grant number TMA2020CDF-3159). The views and opinions of the author expressed herein do not necessarily state or reflect those of EDCTP.

We are equally grateful for the support of the NIH Common Fund, through the OD/Office of Strategic Coordination (OSC) and the Fogarty International Center (FIC), NIH award number U2RTW010672. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the supporting offices.

References

1. WHO guiding principles for pathogen genome data sharing. [Reference Source](#)
2. Aborode AT, et al.: Impact of poor disease surveillance system on COVID-19 response in africa: Time to rethink and rebuild. *Clin Epidemiol Glob Health*. 2021; **12**: 100841. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Makoni M: Africa's \$100-million Pathogen Genomics Initiative. *Lancet Microbe*. 2020; **1**: e318. [PubMed Abstract](#) | [Publisher Full Text](#)
4. Sartorius B, et al.: The burden of bacterial antimicrobial resistance in the WHO African region in 2019: a cross-country systematic analysis. *Lancet Glob. Health*. 2023.
5. Shu Y, McCauley J: GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance*. 2017; **22**. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Risk Ranking and Prioritization of Epidemic-Prone Diseases. *Africa CDC*. [Reference Source](#)
7. Brito AF, et al.: Global disparities in SARS-CoV-2 genomic surveillance. *Nat Commun*. 2022; **13**: 7003. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
8. Tegally H, et al.: The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance. *Science*. 2022; **378**: eabq5358. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Katz K, et al.: The Sequence Read Archive: a decade more of explosive growth. *Nucleic Acids Res*. 2022; **50**: D387–D390. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Mboowa G, Sserwadda I, Aruhomukama D: Genomics and bioinformatics capacity in Africa: no continent is left behind. *Genome*. 2021; **64**: 503–513. [PubMed Abstract](#) | [Publisher Full Text](#)
11. Mulder N, et al.: Genomic Research Data Generation, Analysis and Sharing – Challenges in the African Setting. *Data Sci J*. 2017; **16**: 49. [Publisher Full Text](#)
12. Wilkinson MD, et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016; **3**: 160018. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Various pathogens have been sequenced by institutions within and outside Africa, with the data accessible through the Sequence Read Archive (SRA) of the NCBI from multiple African countries. *figshare*. 2024. [Publisher Full Text](#)
14. The New Public Health Order: Africa's health security Agenda. *Africa CDC*. [Reference Source](#)
15. Nanopore store: Devices. [Reference Source](#)
16. Nanopore store: Price List. [Reference Source](#)
17. Onywera H, et al.: Boosting pathogen genomics and bioinformatics workforce in Africa. *Lancet Infect Dis*. 2023; **24**(2): e106–e112. [PubMed Abstract](#) | [Publisher Full Text](#)
18. Jjingo D, et al.: Bioinformatics mentorship in a resource limited setting. *Brief Bioinform*. 2022; **23**: bbab399.
19. Croxton T, et al.: Building blocks for better biorepositories in Africa. *Genome Med*. 2023; **15**: 92.
20. Pronyk PM, et al.: Advancing pathogen genomics in resource-limited settings. *Cell Genomics*. 2023; **3**: 100443. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Bani Baker Q, et al.: Comprehensive comparison of cloud-based NGS data analysis and alignment tools. *Inform Med Unlocked*. 2020; **18**: 100296. [Publisher Full Text](#)
22. Van de Vuurst P, Escobar LE: Climate change and infectious disease: a review of evidence and research trends. *Infect Dis Poverty*. 2023; **12**: 1–10.
23. Mora C, et al.: Over half of known human pathogenic diseases can be aggravated by climate change. *Nat Clim Chang*. 2022; **12**: 869–875. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Nakkazi E: Cholera outbreak in Africa. *Lancet Infect Dis*. 2023; **23**: 411. [Publisher Full Text](#)
25. de Vries J, et al.: Regulation of genomic and biobanking research in Africa: a content analysis of ethics guidelines, policies and procedures from 22 African countries. *BMC Med Ethics*. 2017; **18**: 8. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Establishment of a Biobanking Network as a Sustainable Mechanism to Accelerate Development and Evaluation of Diagnostic Tests in Africa. *Africa CDC*. [Reference Source](#)
27. Abimiku A, et al.: H3Africa Biorepository Program: Supporting Genomics Research on African Populations by Sharing High-Quality Biospecimens. *Biopreserv Biobank*. 2017; **15**: 99–102. [Publisher Full Text](#)
28. Staunton C, de Vries J: The governance of genomic biobank research in Africa: reframing the regulatory tilt. *J Law Biosci*. 2020; **7**: lsz018. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Klingstrom T, Bongcam-Rudloff E, Reichel J: Legal & ethical compliance when sharing biospecimen. *Brief Funct Genomics*. 2018; **17**: 1–7. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Considerations for developing a national genomic surveillance strategy or action plan for pathogens with pandemic and epidemic potential. [Reference Source](#)
31. Cutler DM, Summers LH: The COVID-19 Pandemic and the \$16 Trillion Virus. *JAMA*. 2020; **324**: 1495–1496. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. The Pandemic Fund Announces First Round of Funding to Help Countries Build Resilience to Future Pandemics. *World Bank*. [Reference Source](#)
33. A Breakthrough for the African Vaccine Manufacturing. *Africa CDC*. [Reference Source](#)
34. Saied AA: Building Africa's first mRNA vaccine facility. *Lancet*. 2023; **402**: 287–288. [PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:    

Version 2

Reviewer Report 28 October 2024

<https://doi.org/10.5256/f1000research.170999.r330025>

© 2024 Folarin O. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Onikepe Folarin

Redeemer's University, Ede, Osun, Nigeria

The authors review is the genomics efforts particularly for disease diagnosis in Africa. The authors made reference to the COVID-19 pandemic and subsequent efforts in Africa to develop capacity in genomics in order to effectively control and manage the menace of infectious diseases which is a major public health challenge in Africa. In addition, the authors touched on challenges facing genomics research and capacity in Africa which include capital, human capacity, supply chain etc. The review showed the fact that Africa is making efforts to get on the fast moving genomics train however with the challenges stated it might take a while to get Africa fully on board.

One of the the statement i did not see in the review is the fact that while Africa is able to sequence lots of the COVID-19 cases, this was done by the efforts of some genomics centers/Institutions thus the more need for the capacity (both human and infrastructure).

While training is emphasized in the recommendations, other contributions to be considered include African government buying in to genomics and genomics research. The same way Africa decided to take charge of controlling COVID-19 during the pandemic, the African government need to support and invest in genomics in Africa. This by allowing the establishment of biotechnology in Africa to take care of challenges of supply chain as well as support genomics research in the country. Another suggested recommendation is the need for Africans to collaborate in genomics research which also was one of the issue that played out during COVID-19 where African scientists collaborated to share methodologies and facilities.

The manuscript is well written and content is adequate for indexing considering other points raised.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Infectious diseases; Genomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 25 October 2024

<https://doi.org/10.5256/f1000research.170999.r330026>

© 2024 Wagner G. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Gabriel E. Wagner 

Medical University of Graz, Graz, Austria

The opinion article, "*The rise of pathogen genomics in Africa*" by Gerald Mboowa and colleagues provides an detailed overview of the sequencing capacities and the efforts taken to further establish molecular pathogen surveillance in Africa during the COVID-19 pandemic.

The article provides a brief introduction to genomic pathogen surveillance and why it is so important to build capacity for the necessary sequencing equipment and bioinformatics. Mboowa and colleagues show very nicely what has happened in this area during the COVID-19 pandemic and how this considerable progress has led to important contributions in and from Africa. There is no question that these competences should continue to be maintained and expanded, as this benefits not only Africa but all of us as a global scientific community - as very well demonstrated on the example of SARS-CoV-2 in the publication. One of the major factors in this area, as the article argues, will definitely be to offer comprehensive bioinformatics expertise and training - the initiatives mentioned in the article can serve as a starting point for the interested reader. The article is definitely an important contribution and provides a great overview of progress and problems in pathogen surveillance in Africa, although a few small changes could further improve readability and impact.

- In the abstract: „ However, the majority of high-income countries sequenced >0.5% of their COVID-19 cases, unlike low- and middle-income countries.“. I would add a number here, what is the difference to low and middle income countries.
Of note, also for the paragraph in the text: I am not convinced that the percentage of sequenced COVID cases is a good estimator. It depends on which COVID tests were used and how extensively they are used. Furthermore, one could argue that waste water surveillance also provides a very good overview and is very cost-efficient.
- In my view, the publication would benefit from including examples of what genomic

surveillance has already achieved in Africa. It is very well elaborated what capacities exist and what needs to be established on a larger scale, but what is the benefit when all goals are achieved? This would emphasize the importance of the topic in my opinion, e.g. extend on the link between NGS and vaccine adaptations.

- In general, the article is very "COVID-heavy". Based on the title, I would have expected a somewhat broader scope here. Could other initiatives for other pathogens perhaps be mentioned here?

In this context. The authors mention: 'The continent still has several pathogens of epidemic and pandemic potential that require near real-time tracking', a brief overview of which pathogens are meant or what 'near real-time tracking' is used for would be useful here. It is not entirely clear which time period is meant.

- Table 1 is pretty big and it's certainly important for networking purposes so it should not be excluded. However, I might suggest to move the country list with the specific NGS equipment to the supplementary.
- The key applications in Table 1 are a bit misleading in my opinion, e.g. one could easily run 12-20 bacterial WGS samples on a minion (or even more on a MiSeq). It all depends on the number of samples and the amount of data required for the analysis. It may be better to have two tables here. One that gives information about the devices and their output and a second one with applications and their requirements? Depending on the number of samples one can then look for the right device here.
- The information on the MinION is not consistent, in Table 1 an output of 50Gb is specified, on page 9, however, 15-20Gb are stated. Further note: The read length for ONT devices is limited mostly by the DNA Prep, the maximum read length is questionable here (especially since different maximum read lengths were specified for different ONT devices). The longest read ever was over 4Mb until recently.
- The search term "Pathogen [All Fields] AND Country [All Fields]" on the SRA webpage (<https://www.ncbi.nlm.nih.gov/sra/>) as describe on page 9 delivers pathogens from all countries, not just Africa?
- The paragraphs under "Major emerging challenges" are a bit redundant and I'd generally suggest to shorten them a bit. In my view, the key messages are sometimes a little lost. The authors could further comment critically here that with cloud platforms, to a certain extent, one also gives the data out of ones hands and makes oneself dependent on certain providers again.
- The paragraph „ Challenges of integrating ethics and law in international collaborations" might also be shorten a bit.

In my view, the points above would emphasize the intentions of the article a little more and increase the flow of reading. Nevertheless, it is already an excellent and above all important article.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular microbiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 25 September 2024

<https://doi.org/10.5256/f1000research.170999.r320742>

© 2024 Bongcam-Rudloff E. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Erik Bongcam-Rudloff 

Department of Animal Biosciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

The revisions made by the authors are satisfactory and align well with the recommendations provided during the initial review.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Partly

Are all factual statements correct and adequately supported by citations?

Partly

Are arguments sufficiently supported by evidence from the published literature?

Partly

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics, genomics, cell biology, molecular biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 03 July 2024

<https://doi.org/10.5256/f1000research.161264.r276910>

© 2024 Gonzalez-Reiche A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Ana S Gonzalez-Reiche**

Icahn School of Medicine at Mount Sinai, New York, USA

The article by Mboowa and collaborators describes the resources available in Africa for genomic pathogen surveillance and how these were built, strengthened, or expanded across the continent due to the COVID-19 pandemic. This important topic deserves attention as it prepares us to respond to emerging infectious diseases globally. Most importantly, it underscores the availability of local capacity and the potential to reduce the gap in research led by African scientists through their perspective.

The article would benefit from highlighting the data generated and the resources available in Africa—for example, Table 1. Rather than listing common NGS platforms, list those available in Africa alone. The table header makes it unclear whether this table is already specific to Africa. Similarly, in Figure 2, the authors could include an insert plot showing the distribution of SRA data for Africa only.

Another essential information would be how these resources are distributed across academic, healthcare, government institutions, and the public and private sectors. This could help lay out a strategic plan to synergize better current efforts to build a sustainable long-term surveillance program. If any of this information is available to the authors, it should be included.

The section "**Major emerging challenges**" could more effectively state the needs if concrete estimates were included. For example, in the third paragraph of this section, the authors could provide an estimate of resources (e.g. computing capacity) and money (e.g. estimated cost per processor/computing node) that would be needed to sustain a genomic surveillance lab sequencing X samples per year. This would be a more practical comparison highlighting the increased cost of sequencing in low-income countries where the supply chains for sequencing reagents aren't as robust as those available to high-income countries. Using one or two of the author's institutions as examples should suffice.

Similarly, in paragraph 6, it seems to this reviewer that investing in building local capacity for bioinformatic analysis will probably pay off in the long run compared to investing in license-based software. This is important as it will highlight the need to establish a critical mass of bioinformatics expertise, including bioinformatic training at universities, etc. (or support for existing training options). A brief cost analysis could also help to support this.

In the section "**Beyond the Post-COVID-19 genomics...**" The Pandemic Fund initiative should be introduced at its first mention. The authors introduce the Pandemic Fund in the second paragraph, but they first mention it at the end of the first paragraph. Please consider reorganizing this section to address this.

Is the topic of the opinion article discussed accurately in the context of the current

literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Partly

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Partly

Competing Interests: No competing interests were disclosed.**Reviewer Expertise:** Genomic Pathogen Surveillance, SARS-CoV-2, Respiratory Viruses, Viral Sequencing, Bioinformatics**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Reviewer Report 11 June 2024

<https://doi.org/10.5256/f1000research.161264.r278094>

© 2024 Bongcam-Rudloff E. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Erik Bongcam-Rudloff**

Department of Animal Biosciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

The authors discuss an important issue regarding the significant threat that new pathogens pose to humanity. Specifically, the article addresses the situation in the African continent and the need for a Pan-African developed early surveillance system.

The authors accurately describe the use of NGS technologies and the necessity of high-throughput computing systems and high-speed internet to share and analyse common sequence data.

However, the article lacks a more detailed discussion about the need for a critical mass of researchers with the bioinformatics competence to perform these tasks at a pan-African level. The article should address and mention some of the existing education initiatives (e.g., AI4D Africa, H3Abionet, EANBiT, University programmes, etc.) and the imperative need to educate more African scientists in using Artificial Intelligence, HPC systems, NGS sequencing, and bioinformatics analysis.

International collaboration relies on the integration of ethics and law across national legal

systems. The article would benefit from discussing the challenges and problems that can arise when considering ethical and legal issues in such collaborations.

References

1. de Vries J, Munung SN, Matimba A, McCurdy S, et al.: Regulation of genomic and biobanking research in Africa: a content analysis of ethics guidelines, policies and procedures from 22 African countries. *BMC Med Ethics*. 2017; **18** (1): 8 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Klingstrom T, Bongcam-Rudloff E, Reichel J: Legal & ethical compliance when sharing biospecimen. *Brief Funct Genomics*. 2018; **17** (1): 1-7 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Partly

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Medical Sciences, Bioinformatics, Biobanking, Exposome research, Metagenomics, Genomics, HPC

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research