Assessment of the pathogen genomics landscape highlights disparities and challenges for effective AMR Surveillance and outbreak response in the East African community

Julien A. Nguinkal[1*], Yedomon A. B. Zoclanclounon[2], Andrea Molina[1,10], Abdi Roba[3], Ndia M. Nyakio[3], Peter N. Lokamar[3], Néhémie Nzoyikorera[4], Théogène Ihorimbere[4], Joseph Nyandwi[4], Mamdouh A. Aguer[5], James A. Maror[5], Michael Lasuba Lokore[5], Monica Fredrick Francis[6], Lawrence A. Mapunda[6], Medard Beyanga[6], Tonny Muyigi[7], Godfrey Pimundu[7], Susan N. Nabadda[7], Emmanuel Kabalisa[8], Jeanne d'Arc Umuringa[8], Isabelle Mukaga Tare[8], Hakim I. Lagu[9], Emmanuel Achol[9], Jürgen May[1], Muna Affara[1] and Florian Gehre[1]

Abstract

The East African Community (EAC) grapples with many challenges in tackling infectious disease threats and antimicrobial resistance (AMR), underscoring the importance of regional and robust pathogen genomics capacities. However, a significant disparity exists among EAC Partner States in harnessing bacterial pathogen sequencing and data analysis capabilities for effective AMR surveillance and outbreak response. This study assesses the current landscape and challenges associated with pathogen next-generation sequencing (NGS) within EAC, explicitly focusing on World Health Organization (WHO) AMR-priority pathogens. The assessment adopts a comprehensive approach, integrating a questionnaire-based survey amongst National Public Health Laboratories (NPHLs) with an analysis of publicly available metadata on bacterial pathogens isolated in the EAC countries. In addition to the heavy reliance on third-party organizations for bacterial NGS, the findings reveal a significant disparity among EAC member States in leveraging bacterial pathogen sequencing and data analysis. Approximately 97% ($n = 4,462$) of publicly available high-quality bacterial genome assemblies of samples collected in the EAC were processed and analyzed by external organizations, mainly in Europe and North America. Tanzania led in-country sequencing efforts, followed by Kenya and Uganda. The other EAC countries had no publicly available samples or had all their samples sequenced and analyzed outside the region. Insufficient local NGS sequencing facilities, limited bioinformatics expertise, lack of adequate computing resources, and inadequate data-sharing mechanisms are among the most pressing challenges that hinder the EAC's NPHLs from effectively leveraging pathogen genomics data. These insights emphasized the need to strengthen microbial pathogen sequencing and data analysis

*Correspondence:
Julien A. Nguinkal
julien.nguinkal@bnitm.de

Full list of author information is available at the end of the article

capabilities within the EAC to empower these laboratories to conduct pathogen sequencing and data analysis independently. Substantial investments in equipment, technology, and capacity-building initiatives are crucial for supporting regional preparedness against infectious disease outbreaks and mitigating the impact of AMR burden. In addition, collaborative efforts should be developed to narrow the gap, remedy regional imbalances, and harmonize NGS data standards. Supporting regional collaboration, strengthening in-country genomics capabilities, and investing in long-term training programs will ultimately improve pathogen data generation and foster a robust NGS-driven AMR surveillance and outbreak response in the EAC, thereby supporting global health initiatives.

Keywords Public health surveillance, Outbreak response, Capacity-building, Pathogens data sharing, East African community, WHO priority pathogens, NGS capabilities, Pathogen genomics

Background

Bacterial pathogens cause infectious diseases that wreak havoc on Sub-Saharan African public health systems [1, 2]. Timely identification of these pathogens using next-generation sequencing (NGS) approaches is critical for effective disease surveillance, outbreak investigation, and the implementation of robust control measures [3–5]. Conventional methods, while valuable, frequently lack the resolution required for precise strain identification, transmission tracing, and understanding of the underlying genetic mechanisms, particularly in the context of emerging antimicrobial resistance (AMR) [5, 6]. The latest advancements in NGS technologies have paved the way for unprecedented pathogen surveillance, molecular epidemiology, and targeted public health interventions. Pathogen whole-genome sequencing (WGS) has been established as a powerful tool for unraveling the genomic landscape of microbial pathogens, providing intricate insights into their genetic makeup, evolution, and diversity [7–10]. This opens new avenues for detecting and responding quickly and effectively to disease threats and outbreaks [11–13].

The East African Community (EAC) is a regional intergovernmental organization that brings together eight East African countries, including Burundi, Kenya, Rwanda, Tanzania, Uganda, the Democratic Republic of the Congo (DRC), Somalia and South Sudan [14]. The EAC partner States, which have a combined population of approximately 285 million people, face a unique set of public health challenges related to infectious diseases. These challenges include high population density, inadequate or limited healthcare infrastructures, and various ecological and environmental factors that influence infectious disease dynamics [15, 16]. In this context, leveraging the full potential of pathogen genomics tools in public health is critical for effective disease management and control. As part of its commitment to timely and effectively addressing public health concerns in the region, the EAC is actively involved in several ongoing continental and global efforts to integrate and use genomics tools in public health. This includes strengthening National Public Health Laboratories (NPHLs) by facilitating the acquisition of state-of-the-art laboratory equipment, promoting training and capacity-building programs, and establishing a rapid response mobile laboratory network (RRMLN) equipped with molecular and NGS facilities [17–19].

Despite these promising efforts, substantial disparities persist among EAC Partner States in their capabilities to utilize NGS technologies to inform AMR monitoring and prevention strategies. These disparities are most evident in the limited ability to conduct in-country bacterial pathogen sequencing and bioinformatic analysis, particularly for the Global Antimicrobial Resistance and Surveillance System (GLASS) priority pathogens outlined in the World Health Organization's (WHO) comprehensive global Genomic Surveillance strategy [20]. These pathogens include *Escherichia coli*, *Klebsiella pneumoniae*, *Staphylococcus aureus, Streptococcus pneumoniae, Salmonella spp., Acinetobacter spp., Shigella spp.*, and *Neisseria gonorrhoeae* [20]. National, regional, and global surveillance reports highlight the alarming prevalence of AMR in pathogenic bacteria within African countries, including the EAC Partner States [21, 22]. This poses a significant threat to public health, necessitating the adoption of modern AMR surveillance tools like pathogen NGS. South Africa, for example, stands out as a pioneer on the continent with a well-developed pathogen genomics infrastructure. Their successful application of these capabilities to improve public health surveillance and preparedness serves as a model for other African countries [23]. In contrast, EAC member States face diverse constraints in establishing, operating and sustaining in-country pathogen NGS capacities.

This study assesses the pathogen genome sequencing capabilities of the EAC by analyzing genome assembly metadata from bacterial isolates, focusing mainly on GLASS AMR-priority pathogens. Our objectives are two-fold: first, to evaluate the current sequencing capabilities within the EAC and identify challenges these countries encounter in integrating NGS pathogen data into public health systems; second, to benchmark these capabilities against South Africa's established pathogen NGS infrastructure. We document the region's pathogen NGS output and efforts by analyzing sequence metadata such as the isolation countries, sequencing platforms

used, and the roles of submitting organizations or laboratories. Insights from a questionnaire-based survey helped pinpoint existing gaps and challenges while identifying opportunities for enhancement and further development. This study underscores the urgent need for substantial investments and commitments from both governmental and international partners to bolster the EAC's bioinformatics and pathogen genomic capabilities. We also provide specific recommendations to narrow these gaps and reduce disparities between EAC Partner States and advanced Northern countries, emphasizing the importance of establishing sustainable, local and in-country capacities.

Bacterial NGS data assessment in EAC

Data collection and wrangling

Data on bacterial genome assemblies were obtained from international sequence repositories, including GenBank, BioSamples, and the Bacterial and Viral Bioinformatics Resource Center (BVBRC). The repositories were queried on different days due to the time required for retrieving and downloading metadata and genome sequences. Access occurred between March 30, 2023, and April 10, 2023. These databases provide pathogen sequence data with contextual information about samples and experimental settings. We investigated the metadata attributes of isolation countries, sequencing platforms, genome quality, submission date, and the role of submitting organizations or laboratories.
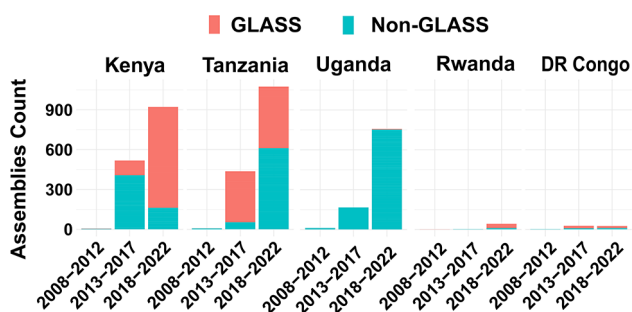
The data retrieved from various sources was integrated while merging disparate records to eliminate redundancy and streamline the dataset into a coherent format. We then performed preprocessing of the data, including cleaning and filtering, to ensure the integrity of the relevant variables. Preprocessing metadata involved removing inconsistent entries between databases and discarding records with missing information, such as

Fig. 1 Temporal trends in the evolution of publicly available bacterial sequences, including AMR-GLASS priority pathogens and other non-GLASS bacteria, for each EAC partner state over the last 15 years, beginning in 2008. Only countries with at least five sequences are displayed for clarity. Burundi and South Sudan are not depicted, as Burundi has two publicly available sequences of non-GLASS pathogens, and South Sudan currently has no publicly accessible pathogen sequences data

isolation countries, sequencing centers, and data owners. Only plasmid and genome sequences flagged as "poor" quality were removed from genomic data. These steps were necessary to ensure the reliability and consistency of downstream analyses. We limited our study only to bacterial isolates from the EAC countries. Finally, we limited the temporal scope to the last 15 years, dating back to 2008. This temporal constraint ensured our analysis remained contemporary and reflected recent advances in NGS technologies.
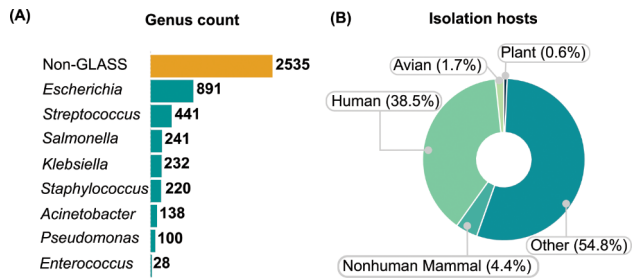
We conducted an online questionnaire-based survey between October 30 and November 14, 2023, to assess the NGS capabilities and challenges in six NPHLs in the EAC. These NPHLs included Burundi, Rwanda, Tanzania, Kenya, Uganda, and South Sudan. The questionnaire, which included closed-ended and open-ended questions, aimed to collect quantitative data and qualitative insights into various aspects of pathogen NGS capabilities and implementation. These aspects included infrastructure, expertise, workflows, challenges, and perspectives or expectations for pathogen NGS implementation in their laboratories. The full questionnaire and raw responses are available in Supplement Files 1 and 2.

Data analysis

Over the past five years, the number of publicly accessible bacterial isolate sequences from the EAC region has substantially increased. Starting with only 700 datasets in 2018, the number of complete bacterial genome sequences has risen to over 4,600 high-quality bacterial sequences by 2022 (Fig. 1). Tanzania leads in terms of overall bacterial genome sequences ($n=3,267$), with 68% ($n=2,223$) representing GLASS pathogens proportion. Kenya follows, with 2,863 sequences, of which 50.4% ($n=1,443$) are AMR-GLASS pathogens. Despite a lower sequence count ($n=145$), Rwanda has a GLASS sequencing rate of 31% ($n=45$). In contrast, Burundi reported only two bacterial genomes, whereas South Sudan lacked publicly available data. Approximately 43.5% of all EAC-isolated samples were GLASS-priority pathogens, and 38.5% of these pathogens had humans as isolation hosts. These pathogens were predominantly *Escherichia coli*, *Streptococcus spp.*, and *Salmonella spp.*, with 891, 441, and 232 sequences, respectively (Fig. 2). When considering in-country sequenced GLASS-priority pathogen samples, Tanzania ($n=134$; 6%) led, followed by Kenya ($n=44$; 3%) and Uganda ($n=4$; 0.5%). When assessing unsubmitted data in the NPHLs, Kenya had the most bacterial sequences ($n=546$), followed by Uganda ($n=85$) and Tanzania ($n=10$), while the other countries reported no unsubmitted sequences (Table 1). We also noticed that no EAC countries had sequence data on *Shigella species* and *Neisseria gonorrhoea*, which is concerning as

(A)   **Genus count**   **(B)**   **Isolation hosts**

Fig. 2 Summary bacterial genome sequences by species and isolation hosts. Cumulative count of bacterial genome sequences at the genus level, including GLASS and non-GLASS pathogens (**A**); Isolation hosts of GLASS-Priority pathogens (**B**)

these species are among the high-priority GLASS-AMR pathogens monitored by the WHO.

### Antimicrobial profile analysis

To identify the AMR determinants associated with these bacterial sequences, we utilized two well-known tools, AMRFinder [23] and StarAMR [24], for predicting genotypic AMR profiles in the bacterial dataset. These tools analyze genomic data to identify known AMR genes (ARGs) and predict resistance phenotypes to specific antimicrobial agents. Additionally, we conducted a longitudinal analysis to examine the prevalence of resistance determinants in the EAC region over the past five years. The data indicate that more than 75% of isolates exhibited resistance to at least one drug from the classes of cephalosporin, tetracycline, beta-lactam, aminoglycoside, folate inhibitors, and fluoroquinolone (Fig. 3). Several ARGs demonstrate high prevalence across multiple bacterial species. Notably, the *gyrA_D87D* gene shows a substantial frequency among *Salmonella* isolates. However, certain genes appear to be specific to particular bacterial

Table 1 Summary of total number of AMR-GLASS pathogens sampled in processed by EAC's NPHLs between 2018 and 2023

Species	Kenya	Burundi	Rwanda	Uganda	Tanzania	South Sudan	DR Congo
Acinetobacter spp.	13	0	0	0	0	0	0
Escherichia coli	346	0	0	0	0	0	0
Klebsiella pneumoniae	34	0	0	0	10	0	0
Neisseria gonorrhoeae	0	0	0	0	0	0	0
Salmonella spp.	153	0	0	65	0	0	0
Shigella spp.	0	0	0	0	0	0	0
Staphylococcus aureus	0	0	0	0	0	0	0
Streptococcus pneumoniae	0	0	0	0	0	0	0
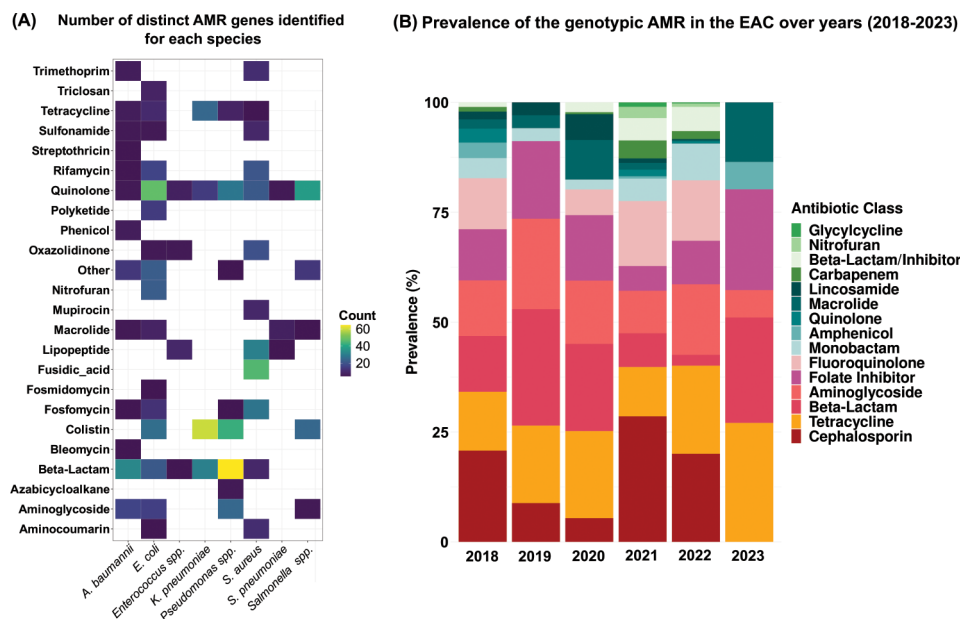
Fig. 3 Antibiotic resistance profiles and key AMR genes in bacterial isolates in the EAC region. (**A**) summarizes the prevalence of predicted resistance genes. This heatmap shows the number of resistance genes identified for different bacterial species and antibiotic classes in the East African Community (EAC) region. Dark purple color indicates a lower number of genes identified. White boxes indicate no resistance was predicted for that specific antibiotic class. (**B**) shows temporal trends in antibiotics resistance patterns to different antibiotic classes. The stacked bar graph depicts the overall prevalence of predicted genotypic resistance profiles in the EAC region for each year between 2018 and 2023
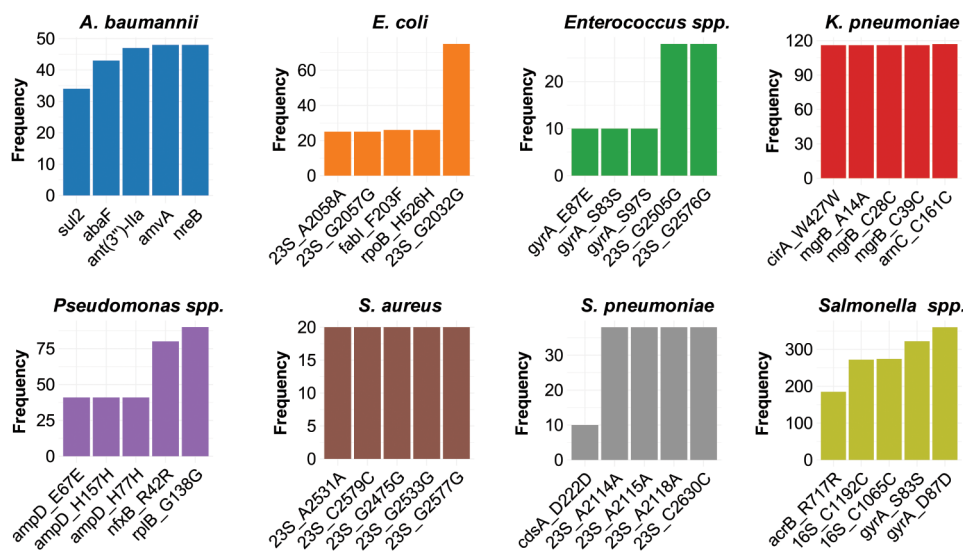
Fig. 4 Most prevalent antibiotic resistance genes identified in bacteria isolates in the EAC region. The bar graphs depict the frequency of the top 5 predicted antibiotic resistance genes (ARGs) identified for each bacterial species isolated in the East African region. Known ARGs were predicted on whole genome sequences using both AMRFinder and STARAMR pipelines
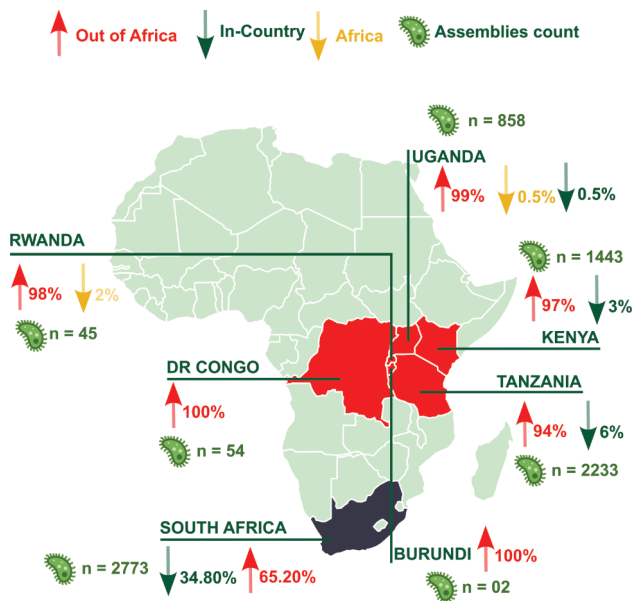
Fig. 5 Imbalance of sequencing efforts for AMR-priority pathogens in the EAC. The map highlights a critical imbalance between isolation countries and organizations performing sequencing and analysis of AMR-GLASS priority pathogens in the East African Community (EAC) Partner States compared to South Africa, a leader on the continent in leveraging genomic and bioinformatics for public health surveillance and preparedness. Red arrow: Highlights the significant proportion of isolates sequenced by organizations outside Africa, indicating reliance on external resources. Green arrow: Indicates a smaller, but crucial, contribution from In-Country organizations to sequencing efforts. Yellow arrow: Shows the contribution from other African countries, suggesting potential regional collaboration opportunities between the EAC region and other African countries

species. For instance, the *arnC_C161C* gene was predominantly found in *K. pneumoniae*, while the resistance gene *23S_G2032G* was more prevalent in *E. coli* (Fig. 4, Supplement file 3).

Challenges and opportunities of pathogens genomics

Current challenges in NGS pathogen analysis

Despite the enormous potential of pathogen NGS in addressing public health issues, EAC Partner States have yet to fully capitalize on these benefits because of various challenges. The survey assessing NPHLs' NGS and data analysis capabilities identified challenges such as limited data storage capacity, a shortage of reagents, and a need for more trained staff in NGS data analysis, including pipelines and workflow deployment. All NPHLs express urgent needs for high-performance computing infrastructure, enhanced training programs, provision of computational resources and equipment, improved internet connectivity, and dedicated bioinformatics workspaces. These constraints extend to various facets, including under-equipped laboratory infrastructure, a scarcity of bioinformatics expertise, insufficient data management capabilities, and a pressing need for standardized protocols and stringent quality control measures [25, 26]. Another major challenge is relying heavily on external organizations and resources for regional NGS analysis. Organizations outside the EAC and Africa perform sequencing and data analysis of approximately 97% (*n*=4,462) of the publicly available bacterial pathogen data (Fig. 5; Table 2) collected in the EAC region. This reliance on external contributors limits the EAC's direct control over data generation, analysis, and long-term

Table 2 Summary of top Organizations/Institutions sequencing AMR-GLASS pathogens sampled in the EAC

Sequencing Organization/Institution	Region	Country	Total_Genomes
Wellcome Trust Sanger Institute	Europe	United Kingdom	617
University of Wisconsin - Madison	North America	USA	499
University of South Carolina	North America	USA	483
University of Maryland	North America	USA	379
University of Queensland	Oceania	Australia	329
Washington State University	North America	USA	276
Broad Institute	North America	USA	101
Kilimanjaro Clinical Research Institute	EAC	Tanzania	76
DOE Joint Genome Institute	North America	USA	62
U.S. Army MRD–Africa	North America	USA	61
University of Bergen	Europe	Norway	38
University of Bern	Europe	Switzerland	36
University of Saskatchewan	North America	Canada	35
Massey University	Oceania	New Zealand	33
Max Planck Institute for Terrestrial Microbiology	Europe	Germany	28
University of Greifswald, Institute of Pharmacy	Europe	Germany	25
J. Craig Venter Institute (JCVI)	North America	USA	20
University of Zurich	Europe	Switzerland	20
Swiss Tropical and Public Health Institute	Europe	Switzerland	19
Los Alamos National Laboratory	North America	USA	17
Uniformed Services University of the Health Sciences	North America	USA	17
University of Copenhagen	Europe	Denmark	17
Masinde Muliro University of Science and Technology	EAC	Kenya	15

interpretation. However, for effective surveillance, it is critical to ensure that genomic data is shared and analyzed promptly. Moreover, ethical and data privacy concerns may constrain the successful implementation of NGS in public health. Overcoming these challenges is, therefore, essential for integrating systematic genomic surveillance into national infection control programs.

Relevance of the NGS-driven public health intervention

Developing and expanding local expertise in NGS analysis is critical for several reasons. First, understanding regional pathogen diversity is crucial for effective surveillance and outbreak response. With pathogen NGS, local public health experts have valuable knowledge about the EAC region's unique microbial diversity, allowing them to quickly identify emerging threats and implement appropriate preparedness and control measures [27]. Second, accurate and timely pathogen identification is essential for monitoring AMR patterns. By analyzing pathogen samples within the region, experts can collect critical data on the prevalence and distribution of resistance genes and inform evidence-based antimicrobial stewardship strategies [28].

Ongoing pathogen genomics efforts in the region

Several African countries, including those within the EAC, recognize the potential of genomics and bioinformatics to transform their public health landscapes. Initiatives such as the Human Heredity and Health in Africa (H3Africa) project [29], the African Society for Laboratory Medicine (ASLM) [30], and the African Centers for Disease Control and Prevention (Africa CDC) have helped to promote research, capacity building, and technology transfer in pathogen genomics and bioinformatics in East Africa.

Within the EAC, individual Partner States have also taken strides towards promoting pathogen genomics and bioinformatics. National research institutions, universities, and health ministries are increasingly collaborating to build local expertise and infrastructure for genome sequencing and bioinformatics analysis. The EAC engages in international partnerships and consortia to pool resources, share knowledge, and harmonize efforts to address common public health challenges. Some of these initiatives are spearheaded and supported by multi-country organizations and consortia such as the Africa Pathogen Genomics Initiative (Africa PGI) [30], SeqAfrica [31], the East African WHO Regional Office, and the Centre for Epidemic Response and Innovation (CERI) [32]. They develop and sustain microbial genomics and bioinformatics capacity within the continent. Their primary focus is to strengthen AMR and infectious disease surveillance by providing necessary instrumentation,

training in equipment usage, and essential bioinformatics skills for data analysis.

At the national levels, since 2017, the EAC Secretariat, in collaboration with the Bernard Nocht Institute for Tropical Medicine (BNITM) and the German Federal Ministry for Economic Cooperation and Development (BMZ) through the German Development Bank Kreditanstalt für Wiederaufbau (KfW), has been executing the EAC Regional Network of Public Health Reference Laboratories for Communicable Diseases and AMR Project across the EAC Partner States. This project includes strengthening NPHLs by facilitating the acquisition of laboratory equipment and establishing an RRMLN equipped with NGS instruments [17–19]. It also involve training national experts according to the 'Training of Trainers' (ToTs) concept. With this, the EAC aims to strengthen the ability of its Partner States to efficiently prepare and respond to infectious disease outbreaks, including cross-border epidemics and priority AMR pathogens.

Strategic enhancement of the regional Bioinformatics expertise and NGS capabilities
(i) Expanding capacity-building and training programs.
Strategic investments must strengthen EAC's pathogen genomic and bioinformatics capacity. Funding agencies and governmental organizations should prioritize infrastructure and training programs that foster local expertise. Training concepts tailored to the region's needs should include hands-on practical sessions to familiarize researchers, clinicians, and laboratory technicians with NGS technologies, data analysis pipelines, and result interpretation. In addition, capacity development programs and establishing local bioinformatics centers equipped with high-performance computing infrastructure could be pivotal for a sustainable bioinformatics infrastructure. These regional genomics centers can be hubs for bioinformatics training, collaborative research, and data analysis services. They can also facilitate the development of standardized protocols and workflows for pathogen NGS analysis, ensuring the highest data quality and comparability across EAC Partner States.

(ii) Collaborative initiatives.
Promoting collaborative pathogen genomics research initiatives between EAC research institutions (NPHLs, academic institutions, Ministries of Health, and non-governmental organizations) and international partner organizations represents a promising avenue for addressing regional challenges in the local analysis of microbial NGS data. Such collaborations, for example, provide many benefits, such as knowledge exchange and access to shared resources unavailable in most EAC countries. They can pave the way for public health biobanks. These biobanks are critical for understanding pathogen genetic variability and susceptibility and supporting healthcare research strategies for AMR and other infectious disease surveillance, prevention, and control.

(iii) Advocacy and policy support.
Governments in the EAC Partner States can influence the region's pathogen genomics and bioinformatics landscape. The EAC Partner States could collaborate to develop a regional genomic surveillance strategy. Countries, in particular, should advocate for establishing regional platforms for the transparent and timely sharing of pathogen genomic data. Such platforms should promote a culture of open data sharing, ensuring valuable information is accessible to NPHLs, health ministries, and healthcare practitioners. In addition, they should also develop legal and regulatory frameworks to ensure ethical and responsible use of pathogen genomics data. Finally, each EAC Partner State must establish a system and infrastructure to ensure harmonization in quality, performance, interpretation, and documentation of NGS data. This will facilitate and support the implementation of an in-country and sustainable pathogen genomics capacity.

Discussion
We have assessed and underscored the significance of pathogen NGS and bioinformatics for public health in the EAC. As highlighted in previous studies and consortium meetings [33, 34], there is a growing recognition of the need to enhance pathogen surveillance capabilities, leverage opportunities for resilient public health systems, and improve global pathogen surveillance and public health preparedness. Our study underscores the heavy dependence on external organizations and resources for sequencing and data analysis of bacterial pathogens isolated within the EAC region, as illustrated in Figs. 5 and 6. This reliance results from the shortage of local expertise and inadequate facilities and capacities for independent NGS data analysis [7]. While collaboration with international partners has been instrumental in implementing pathogen NGS and narrowing gaps with global Northern countries, building and expanding local sequencing, computing, and NGS analysis capacity is essential for a sustainable pathogen genomics and bioinformatics landscape [35]. Relying on external organizations may lead to issues with data ownership, accessibility, and delays in obtaining critical pathogen data generated outside the region. Particularly in the context of infectious disease outbreaks, prompt intervention is crucial, and any delay in sequencing and data analysis might hinder countries' ability to make timely decisions for real-time surveillance. Therefore, establishing local data-sharing and analysis capabilities is a crucial pillar.

Despite these challenges, promising initiatives are underway in the region to pool resources, share
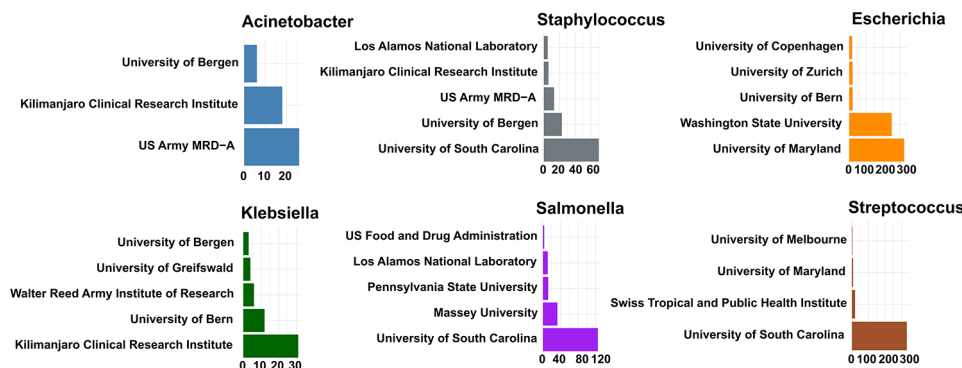
Fig. 6 Top organizations performing NGS and submitting AMR-GLASS priority pathogens in the EAC

knowledge, and harmonize efforts to address common public health challenges. High-throughput sequencing of microbial isolates has gained traction in the EAC region, partly due to global efforts during the COVID-19 pandemic [28, 34, 36]. These regional efforts and increased awareness are gradually reshaping the microbial genomics landscape in the region. However, sustained investments in local capacity-building and resource allocation are needed to achieve NGS self-sufficiency and fully utilize the potential of pathogen genomics to guide effective public health interventions, outbreak preparedness and response. Countries that have consistently invested in NGS technology and laboratory infrastructure have demonstrated rapid national responses to infectious disease threats, as evidenced by their response to the COVID-19 pandemic [8, 10]. A notable example is South Africa, which concentrates nearly 40% of the continent's genomics capacities while demonstrating high pathogen sequencing capabilities [25].

We compared the EAC pathogen data with data from South Africa, a pioneer on the continent with robust pathogen genomic infrastructure and success stories of leveraging these genomic capacities to improve public surveillance and preparedness. This data shows clearly that, due to the deployment of local pathogen genomic and bioinformatics capacity, the in-country-sequencing and data analysis capabilities are substantially better (~35%) in South Africa than the average (~3%) in the EAC region, supporting that the lack of local skill personal and infrastructure is associated with low in-country sample analysis in these countries (Fig. 6). This reinforces the importance of strategic investment in genomic infrastructure, particularly in regions with significant healthcare burdens. South Africa's robust genomic capacity positions it as a leader in pathogen surveillance and outbreak response within Africa and highlights the potential for leveraging such capabilities to address broader global health challenges. Therefore, pathogen genomics and data analysis, such as bioinformatics, need to be mainstreamed locally to benefit communities effectively.

Systematically shipping samples abroad neither supports the regional public health systems nor benefits the host institutions on the continent. Addressing these gaps is crucial by prioritizing sustained regional capacity-building efforts and funding initiatives within Africa and the EAC.

Conclusion

In conclusion, assessing pathogen genomic capabilities in the EAC highlights critical challenges and underscores the urgent need for strategic interventions to improve public health surveillance, intervention, and preparedness. Addressing these challenges requires concerted efforts and a multi-level strategy, including collaborative initiatives involving local universities, research institutions, and international partners. Strengthening local NGS capacity and expertise, establishing regional data-sharing and analysis platforms, and developing regulatory frameworks for pathogen data management are essential steps. Enhanced regional collaboration in setting quality standards for processing, reporting, and interpreting pathogen genomics data can optimize these initiatives. These efforts will empower NPHLs and national laboratories to independently perform pathogen sequencing and data analysis, improving disease surveillance, outbreak preparedness, and response capabilities. This will ultimately support the development of sustainable NGS data-driven national strategic plans to strengthen public health systems in the EAC Partner States.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-024-18990-0.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Declarations

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Infectious Disease Epidemiology, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany
[2]Plant Sciences and the Bioeconomy, Rothamsted Research, Harpenden AL5 2JQ, UK
[3]Department of Disease Surveillance and Epidemic Response, Ministry of Health, Nairobi, Kenya
[4]National Reference Laboratory, National Institute of Public Health, Bujumbura, Burundi
[5]National Public Health Laboratory, Ministry of Health, Juba, Republic of South Sudan
[6]National Public Health Laboratory, Ministry of Health, Dar es Salam, Tanzania
[7]Central Public Health Laboratories, National Health Laboratories, Ministry of Health, Kampala, Uganda
[8]Biomedical Services Department, Biomedical Centre Rwanda, Kigali, Rwanda
[9]Health Department, East African Community (EAC), Arusha, Tanzania
[10]Animal Science School, University of Costa Rica, San José, Costa Rica

References
1. Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, Han C, Bisignano C, Rao P, Wool E, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet. 2022;399(10325):629–55.
2. Ikuta KS, Swetschinski LR, Aguilar GR, Sharara F, Mestrovic T, Gray AP, Weaver ND, Wool EE, Han C, Hayoon AG, et al. Global mortality associated with 33 bacterial pathogens in 2019: a systematic analysis for the global burden of disease study 2019. Lancet. 2022;400(10369):2221–48.
3. Gordon AK, Marais B, Walker TM, Sintchenko V. Clinical and public health utility of mycobacterium tuberculosis whole genome sequencing. Int J Infect Dis. 2021;113:40–2.
4. Armstrong GL, MacCannell DR, Taylor J, Carleton HA, Neuhaus EB, Bradbury RS, Posey JE, Gwinn M. Pathogen genomics in public health. N Engl J Med. 2019;381(26):2569–80.
5. Tang P, Croxen MA, Hasan MR, Hsiao WW, Hoang LM. Infection control in the new age of genomic epidemiology. Am J Infect Control. 2017;45(2):170–9.
6. Ferdinand AS, Kelaher M, Lane CR, Silva AG, Sherry NL, Ballard SA, Andersson P, Hoang T, Denholm JT, Easton M, et al. An implementation science approach to evaluating pathogen whole genome sequencing in public health. Genome Med. 2021;13:1–11.
7. Christoffels A, Mboowa G, Heusden P, Makhubela S, Githinji G, Mwangi S, Onywera H, Nnaemeka N, Amoako DG, Olawoye I, et al. A pan-african pathogen genomics data sharing platform to support disease outbreaks. Nat Med. 2023;29(5):1052–5.
8. Akande OW, Carter LL, Abubakar A, Achilla R, Barakat A, Gumede N, Guseinova A, Inbanathan FY, Kato M, Koua E, et al. Strengthening pathogen genomic surveillance for health emergencies: insights from the world health organization's regional initiatives. Front Public Health. 2023;11:1146730.
9. Grant K, Jenkins C, Arnold C, Green J, Zambon M. Implementing pathogen genomics: a case study. London: Public Health England; 2018.
10. Revez J, Espinosa L, Albiger B, Leitmeyer KC, Struelens MJ, Tóth A, Grískevícius A, Vatopoulos A, Skoczynska A, Pantosti A et al. Survey on the use of whole-genome sequencing for infectious diseases surveillance: rapid expansion of European national capacities, 2015–2016. Front Public Health, 347 (2017).
11. AMR NGHRU. Whole-genome sequencing as part of national and international surveillance programmes for antimicrobial resistance: a roadmap. BMJ Global Health. 2020;5(11):002244.
12. Aanensen DM, Feil EJ, Holden MT, Dordel J, Yeats CA, Fedosejev A, Goater R, Castillo-Ramírez S, Corander J, Colijn C, et al. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive staphylococcus aureus in Europe. MBio. 2016;7(3):10–1128.
13. Brown E, Dessai U, McGarry S, Gerner-Smidt P. Use of whole-genome sequencing for food safety and public health in the United States. Foodborne Pathog Dis. 2019;16(7):441–50.
14. Ogola FO, Njenga GN, Mhando PC, Kiggundu MN. A profile of the East African community. Afr J Manage. 2015;1(4):333–64.
15. Githeko AK. Health-related vulnerabilities and enabling institutions to facilitate responses to climate change in east Africa. East Afr Sci. 2021;3(1):1–18.
16. Azevedo MJ, Azevedo MJ. The state of health system (s) in Africa: challenges and opportunities. Historical perspectives on the state of health and health systems in Africa, volume II: the modern era, 1–73 (2017).
17. Affara M, Lagu HI, Achol E, Karamagi R, Omari N, Ochido G, Kezakarayagwa E, Kabatesi F, Nkeshimana A, Roba A, et al. The east African community (eac) mobile laboratory networks in Kenya, Burundi, Tanzania, Rwanda, Uganda, and South Sudan—from project implementation to outbreak response against dengue, Ebola, covid-19, and epidemic-prone diseases. BMC Med. 2021;19(1):1–15.
18. Gehre F, Lagu H, Achol E, Omari N, Ochido G, Duraffour S, Hinzmann J, Kezakarayagwa E, Kabatesi F, Nkeshimana A, et al. The east African community's mobile laboratory network's rapid response during the first 2 weeks of the Ebola Sudan virus disease (svd) outbreak in Uganda and pandemic preparedness activities in south Sudan, Rwanda, Tanzania, Burundi, Kenya. BMJ Global Health. 2022;7(12):011073.
19. Gehre F, Lagu HI, Achol E, Omari N, Ochido G, Shand K, Alvarado AM, Ruge G, Kezakarayagwa E, Kabatesi F et al. The east African community mobile laboratory network prepares for monkeypox outbreaks. J Public Health Afr (2023).
20. Organization WH. Pathogen Genomics and Precision Medicine: a review of the literature and applications for infectious Disease Control. World Health Organization; 2015. http://apps.who.int/iris/bitstream/handle/10665/193736/9789241509763eng.pdf;jsessionid=F2EBC. 03B879EFDB7F701FF6982F036C1?sequence=1.
21. Ampaire L, Muhindo A, Orikiriza P, Mwanga-Amumpaire J, Boum Y, Bebell L. A review of antimicrobial resistance in East Africa. Afr J Lab Med. 2016;5(1):1–6.
22. Camara N, Moremi N, Mghamba J, Eliakimu E, Shumba E, Ondoa P, Beverly Egyir. Surveillance of antimicrobial resistance in human health in Tanzania: 2016–2021. Afr J Lab Med. 2023;12(1):1–8.
23. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, Hoffmann M et al. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Scientific reports* 11, no. 1 (2021): 12728.
24. Bharat A, Petkau A, Avery BP, Chen JC, Folster JP, Carson CA, Kearney A et al. Correlation between phenotypic and in silico detection of antimicrobial resistance in Salmonella enterica in Canada using Staramr. *Microorganisms* 10, no. 2 (2022): 292.
25. Inzaule SC, Tessema SK, Kebede Y, Ouma AEO, Nkengasong JN. Genomic-informed pathogen surveillance in Africa: opportunities and challenges. Lancet Infect Dis. 2021;21(9):281–9.
26. Lumaka A, Carstens N, Devriendt K, Krause A, Kulohoma B, Kumuthini J, Mubungu G, Mukisa J, Nel M, Olanrewaju TO, et al. Increasing African genomic data generation and sharing to resolve rare and undiagnosed

Nguinkal *et al. BMC Public Health*   (2024) 24:1500

Page 10 of 10

diseases in Africa: a call-to-action by the h3africa rare diseases working group. Orphanet J Rare Dis. 2022;17(1):230.

27. Mwapagha LM. Why pathogen genomics is crucial in Africa's public health. Afr J Lab Med. 2023;12(1):2166.

28. Otu A, Agogo E, Ebenso B. Africa needs more genome sequencing to tackle new variants of sars-cov-2. Nat Med. 2021;27(5):744–5.

29. Mulder N, Abimiku A, Adebamowo SN, Vries J, Matimba A, Olowoyo P, Ramsay M, Skelton M, Stein DJ. H3africa: current perspectives. Pharmacogenomics Personalized Med, 59–66 (2018).

30. ASLM LabCoP [25], Makoni M. Africa's $100-million pathogen genomics initiative. Lancet 11 Microbe. 2020;1(8):318.

31. Carey ME, Dyson ZA, Argimón S, Cerdeira L, Yeats C, Aanensen D, Mboowa G, Baker S, Tessema SK, Smith AM et al. Unlocking the potential of genomic data to inform typhoid fever control policy: Supportive resources for genomic data generation, analysis, and visualization. In: Open Forum Infectious Diseases, vol. 10, pp. 38–46 (2023). Oxford University Press US.

32. Publications by the Centre for Environmental Rights. https://ceri.org.za/publications/. Accessed: July 14, 2023.

33. Meisel JS, Daniel J, Nasko et al. Brian Brubach, Victoria Cepeda-Espinoza, Jessica Chopyk, Héctor Corrada-Bravo, Marcus Fedarko. Current progress and future opportunities in applications of bioinformatics for biodefense and pathogen detection: report from the Winter Mid-Atlantic Microbiome Meetup, College Park, MD, January 10, 2018. *Microbiome* 6 (2018): 1–10.

34. Onywera H, Ondoa P, Nfii F, Ogwell A, Kebede Y, Christoffels A, Tessema SK. Boosting pathogen genomics and bioinformatics workforce in Africa. The Lancet Infectious Diseases; 2023.

35. Alemnji GA, Zeh C, Yao K, Fonjungo PN. Strengthening national health laboratories in sub-S aharan a frica: a decade of remarkable progress. Tropical Med Int Health. 2014;19(4):450–8.

36. Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, Darren P, Martin, et al. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. Science. 2021;374(6566):423–31.

Publisher's Note