Briefings in Bioinformatics, 2026, 27, bbag046
https://doi.org/10.1093/bib/bbag046
Published: 12 February 2026
Review

OXFORD

Pathways, outputs and impact of NIH-supported bioinformatics and genomics graduate trainees in Africa

Daudi Jjingo[1,2,3,*], Andrew Walakira[1], Suhaila Hashim[4,5], Cisse Cheickna[6], Ronald Galiwango[1,2], Caleb Kibet[7], Florence N. Kivunike[8], Gerald Mboowa[1,2,9], Fredrick Elishama Kakembo[1,10], Babajide Ayodele[11], Jean-Baka Domelevo Entfellner[12], Santie de Villiers[4], Karen Wambui[7], Segun Fatumo[13,14], Tinashe Chikowore[15,16], John Mukisa[1,10], Alfred Ssekagiri[1,17], Nicholas Bbosa[1,17], Julius Mulindwa[18], Samuel Kyobe[10], Mike Nsubuga[1,2,19], Grace Kebirungi[1,2], Eric Katagirya[10], Savannah Mwesigwa[10], Ibra Lujumba[1], Rogers Kamulegeya[10], Samuel Kirimunda[10], Stephen Kanyerezi[1,10,20], Shahiid Kiyaga[1,10], Ivan Sserwadda[1,10,20], Davis Kiberu[1,10], Bernard S. Bagaya[2,10], Julius Okwir[1], Patricia Nabisubi[1,10], Grace Nabakooza[21,22], Mugume Twinamatsiko Atwine[1], Ricard Sserunjogi[1,23], Rolanda Julius[24], Mariam Quiñones[25], Meghan McCarthy[25], Phillip Cruz[25], Karlynn Noble[25], Christopher J. Whalen[25,26], Darrell Hurt[25], Maria Y. Giovanni[25], Michael Tartakovsky[25], Deogratius Ssemwanga[14,17], John M. Kitayimbwa[27], Steven J. Reynolds[28,29], Christopher C. Whalen[30], Andrew Kambugu[2], Neil A. Hanchard[31,32], Li Jian[33], Peter Amoako-Yirenkyi[34], Graeme Mardon[31,35], I. King Jordan[36], Samson Pandam Salifu[37], Mamadou Wele[6], Ezekiel Adebiyi[2,11,38], Jeffrey G. Shaffer[39], Seydou Doumbia[40,41], David Patrick Kateete[10], Michelle Skelton[24], Nicola Mulder[24], Jonathan K. Kayondo[42], Daniel Masiga[43], H3Africa Consortium[‡]

[1]The African Center of Excellence in Bioinformatics and Data-Intensive Sciences, Hall Lane, Makerere University, Central, P.O BOX 22418, Kampala, Uganda
[2]The Infectious Diseases Institute, Hall Lane, Makerere University, Central, P.O BOX 22418, Kampala, Uganda
[3]The Department of Computer Science, College of Computing and Information Sciences, University Rd, Makerere University, Kampala, Uganda
[4]Department of Biochemistry and Biotechnology, Kilifi Ganze Rd, Pwani University, Kilifi, Kenya
[5]Pwani University Biosciences Research Centre, Malindi-Garsen Rd, Pwani University, Kilifi, Kenya
[6]The African Center of Excellence in Bioinformatics and Data-Intensive Science, University of Sciences, Techniques and Technologies of Bamako, J287+PM5, Bamako, Mali
[7]Martin Lüscher Emerging Infectious Diseases Laboratory, International Centre of Insect Physiology and Ecology, Ndemi Rd, Nairobi, Kenya
[8]The Department of Information Technology, College of Computing and Information Sciences , Hill Rd, Makerere University, Uganda
[9]Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA 02142, United States
[10]Department of Immunology and Molecular Biology, School of Biomedical Sciences, Upper Mulago Hill Rd, Makerere University College of Health Sciences, P.O. Box 7072, Kampala, Uganda
[11]Covenant University Bioinformatics Research, Department of Computer and Information Sciences, 10 Idiroko Road, Covenant University, Ota, Ogun, 112233, Nigeria
[12]Biosciences Eastern and Central Africa, International Livestock Research Institute Hub, P.O. Box 30709, Nairobi 00100, Kenya
[13]Precision Healthcare University Research Institute, Queen Mary University of London, Mile End Rd, London, United Kingdom
[14]Medical Research Council/Uganda Virus Research Institute and London School of Hygiene and Tropical Medicine Uganda Research Unit, Nakiwogo Rd, PO Box 49, Entebbe, Uganda
[15]Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, 181 Longwood Ave, Boston 02115, United States
[16]Medical Research Council/Wits Developmental Pathways for Health Research Unit, Department of Pediatrics, Faculty of Health Sciences, University of the Witwatersrand, College Rd, Johannesburg, South Africa
[17]Department of General Virology, Uganda Virus Research Institute, Nakiwogo Rd, Entebbe, Uganda
[18]The Department of Biochemistry, College of Natural Sciences, Makerere University, P. O. Box 7062, Kampala, Uganda
[19]Faculty of Health Sciences, University of Bristol, BS8 1QU, Bristol, United Kingdom
[20]National Health Laboratories and Diagnostics Services, Central Public Health Laboratories, Ministry of Health, P.O Box 7272, Kampala, Uganda
[21]Oak Ridge Institute for Science and Education, Oak Ridge, Tennessee, United States, Assigned to Centers for Disease Control and Prevention, P.O. Box 117, Atlanta, GA United States
[22]Eagle Global Scientific LLC, Contracting Agency to the Centers for Disease Control and Prevention, Atlanta, GA United States

---

Received: August 11, 2025. **Revised:** January 6, 2026. **Accepted:** January 16, 2026

[23]AirQo, Department of Computer Science, Makerere University, Plot 56 Pool Rd, Kampala, Uganda
[24]Computational Biology Division, Department of Integrative Biomedical Sciences, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa
[25]Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, United States
[26]Research Data and Communication Technologies Benefit Corp, P.O. Box 67 Garrett Park, MD, United States
[27]Uganda Christian University, Bishop Tucker Rd, Mukono, Uganda
[28]Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD.
[29]Rakai Health Sciences Program, Masaka - Kakuto Rd, Kalisizo, Uganda.
[30]Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, GA, United States.
[31]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, United States.
[32]Childhood Complex Disease Genomics Section, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, United States
[33]School of Public Health and Tropical Medicine, Tulane University, New Orleans, New Orleans, LA 70112, United States
[34]Department of Economics and Statistics, Garden City University College, Kenyasi, Ghana
[35]Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX 77030, United States.
[36]School of Biological Sciences, Georgia Institute of Technology, 10 Ferst Dr NW, Atlanta, GA 30332, United States
[37]Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[38]Applied Bioinformatics Division, German Cancer Research Center (DKFZ), Heidelberg, Baden-Württemberg, 69120, Germany
[39]School of Public Health and Tropical Medicine, Tulane University, 1440 Canal St, New Orleans, LA 70112, United States
[40]Faculty of Medicine and Odontostomatology, University of Sciences Techniques and Technologies of Bamako, M2C3+HW9, Bamako, Mali
[41]University Clinical Research Center, University of Sciences, Techniques and Technologies of Bamako, M2C5+23G, Bamako, Mali
[42]Entomology Division, Uganda Virus Research Institute, 51/59 Nakiwogo Rd, Entebbe, Uganda
[43]International Centre of Insect Physiology and Ecology, Nairobi, Kenya

*Corresponding author. African Center of Excellence in Bioinformatics and Data Intensive Sciences, Hall Lane, Makerere University, Central, P.O BOX 22418, Kampala, Uganda. E-mail: daudi.jjingo@mak.ac.ug

‡The **H3Africa Consortium** is a cross-continent network of researchers and universities funded by the National Institute of Health and Wellcome Trust, with support from the African Society of Human Genetics. It aims to enhance scientific capacity across the continent to investigate the genomics and environmental contributors to health and disease.

Daudi Jjingo is the Director of the African Center of Excellence in Bioinformatics and Data-Int4ensive Sciences, and a Senior Lecturer in the Department of Computer Science at Makerere University in Uganda with research interests in bioinformatics and data science.

Abstract

Global biomedical and health research is increasingly relying on genomic and computational approaches, largely driven by the increasing volumes of nucleic acid sequencing. Concurrently, epidemiological studies and clinical records are generating enormous amounts of data amenable to disease modeling, machine learning, and artificial intelligence techniques. Bioinformatics and data science expertise is therefore essential for improved population health. Accordingly, in 2012, the US National Institutes of Health (NIH) in partnership with the Wellcome Trust, and with support from the African Society for Human Genetics, initiated the H3Africa (Human Heredity and Health in Africa) consortium. One of its key goals was to build capacity among African scientists to lead research on genetic and environmental contributors to health and disease across the continent. In 2017, the NIH provided funding to support the establishment of four graduate bioinformatics training programs across five African universities. Over seven years, these programs enrolled multiple trainees ($n > 270$), with >110 earning Master's degrees and >20 completing PhDs in Bioinformatics. It is thus timely to evaluate the outcomes and impact of these programs, particularly regarding graduation rates, career trajectories, and the institutions and research domains their alumni are serving. We also assess employment outcomes and the nature of the research they are enabling ($n > 110$ peer-reviewed articles). We additionally include the progress and outputs of the programs' instructors, which were partially enabled by program resources, networks, and trainees. Overall, this review paints valuable insights into the pioneering role of NIH extramural support in shaping Africa's biomedical research landscape.

Keywords post-graduation pathways, bioinformatics, genomics, graduate trainees, Africa

Introduction

Africa and low- and middle-income countries (LMICs) in general bear a disproportionately higher burden of killer infectious diseases such as Malaria, leishmaniasis, tuberculosis (TB), and human immunodeficiency virus in terms of both morbidity and mortality. Furthermore, the continent has a high endemicity of neglected parasitic human and animal diseases like schistosomiasis and trypanosomiasis [1–3]. Moreover, we are seeing an increasing incidence of non-communicable diseases (NCDs) like diabetes and cancer on the continent [4, 5], owing to longer life expectancy, urbanization, and changes in lifestyle and diet. Acute shortages of healthcare workers, health facilities, and sub-optimal biomedical research exacerbate these problems. Biomedical research and its translation into public health is progressively becoming dependent on computing and computational platforms (bioinformatics, data science, and artificial intelligence), which portend substantial improvements in efficiencies of health processes, accuracy and speed of research.

Bioinformatics is a branch of data science that involves the acquisition, storage, analysis, interpretation and dissemination of biological and health data, most often molecular sequencing data on macromolecules essential to how life functions like deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and proteins. It sits at the interface between clinical and basic science research, as it draws upon concepts of population science and the knowledge and theory of biological and computational science. Because it is a so-called 'dry' science conducted largely within the computational space, it tends to be the cheaper, quicker and more scalable research component. Indeed, the last twenty years have seen the migration of a significant component of biomedical research from the wet laboratory to the 'dry' computer [6–8], enabled by the simultaneous technological developments in high-throughput sequencing [9, 10] and increased computational power. The relatively cheaper cost and faster execution of bioinformatics as a digital platform for research in comparison to wet laboratory experimentation is particularly appropriate for resource-limited settings since it is consequently more scalable. Furthermore, the increasing potential and corresponding importance of data science and artificial intelligence in biomedical research [11, 12] underscores the importance of building capacity in these data-intensive sciences within Africa to help leap-frog the continent from its relative lag in basic, clinical and biomedical research [13, 14]. Yet, sufficiently skilled bioinformaticians are still few and far between within LMICs, which constitutes a critical gap for the achievement of the scientific promise of bioinformatics, data science and artificial intelligence in health.

In 2017, the National Institutes of Health (NIH), through its Fogarty International Institute (FIC) and funding from the NIH Common Fund, announced five-year support to start four bioinformatics and genomics degree training programs across Africa. These were, respectively: Nurturing Genomics and Bioinformatics Research Capacity in Africa (BRECA) based at Makerere University in Uganda, the East African Network for Bioinformatics Training (EANBiT) based at the International Centre of Insect Physiology and Ecology (ICIPE) in Kenya but coordinating training in Kenya and Uganda, the West African Sustainable Leadership And Innovation Training In Bioinformatics Research (WASLITBRe) hosted at Covenant University (CU) in Nigeria but additionally working with the Kwame Nkrumah University of Science and Technology (KNUST), Ghana, and University of Sciences, Techniques & Technologies of Bamako (USTTB) in Mali and the West African Center of Excellence for Global Health Bioinformatics Research Training (WACE-B) based at USTTB in Mali and working with other African Francophone Countries. Many of these groups were members of H3ABioNet, a pan-African Bioinformatics Network within the H3Africa initiative, which ran short courses in bioinformatics for 12 years and developed guidelines for starting bioinformatics degree programs early in the project [15, 16]. H3ABioNet also contributed to faculty and mentorship for the new degree programs. Although several other bioinformatics trainings, usually in the form of short-term workshops or courses ranging from a few days to a couple of months have been and continue to be carried out on the continent [17, 18], the programs described here stood out in two important ways. First, it was specifically designed for long-term (at least two years) degree training, recognizing as it did the need for in-depth, sustained instruction and mentorship as prerequisites for the creation of a professional, research-competent bioinformatics workforce. Secondly, it was executed as multiple programs, spanning multiple universities across different countries on the continent. It thus took on an unprecedented level of scale in terms of duration of training and geographical scope. Although the programs were largely autonomous, they all received broad oversight from the NIH. They maintained a level of coordination through shared materials, biannual consortium meetings and sometimes shared instructors and curriculum discussions.

Justification for long-term bioinformatics training

Bioinformatics scientists are individuals who can apply bioinformatics and data science principles as independent scientists—with the ability to conceive and answer critical biomedical questions using sound scientific principles with relevant bioinformatics techniques [19]. They can be contrasted with bioinformatics engineers or technicians, who can apply bioinformatics techniques to specified problems but may not be well grounded in biological and biomedical principles to conceive and perform independent research [19]. Skills for the latter can often be gained in short-term training of days, weeks or a few months. Bioinformatics scientists, on the other hand, need to be thoroughly grounded in all major biological, bioinformatics and data science domains [20, 21], including Next Generation Sequencing (NGS), sequence alignment and analysis, phylogenetics, bioinformatics programming (python, R and Unix), primary and secondary bioinformatics databases, molecular biology, population genetics, molecular evolution, and disease modeling. That range of knowledge and skills, together with the required depth, demands sufficient time [18, 20, 21] for learners to digest, contextualize, discuss and practice skills and knowledge in stratified sequential phases that often take years. Such learning is best delivered in longer-term in-person settings, as is the case for high schools [18, 22–25], undergraduate [26, 27], graduate degree, and postdoc programs [28]. It requires years of commitment from both the learners and the instructors. It is thus more susceptible to resource constraints like the paucity of long-term expert trainers and/or mentors, sustained computing infrastructure and effective accreditation bureaucracies. Indeed, this is evidenced by only a few such programs in Africa [28] and elsewhere [29]. The goal and support from the Fogarty International Center and NIH Common Fund [30, 31] were thus to initiate and sustain fully accredited long-term bioinformatics education degree training programs in Africa. Be that as it may, some non-degree training was also supported.

Master of Science versus Doctor of Philosophy training rationale

The programs were designed mostly for MSc bioinformatics training to meet the urgent need for those skills in Africa, and to achieve an initial critical mass of bioinformaticians and data scientists in a reasonable turn-around time (2 years). The choice was also informed by the relatively cheaper cost of MSc training relative to PhD training. Furthermore, PhD training requires a master's degree, so we needed to form a large pool of master's degrees to eventually feed into the PhD programs. The average PhD:MSc cost ratio of ∼3:1 would mean the attainment of at least thrice more trainees [32]. In our cohorts, the ratio of MSc to PhD enrollees was 4:1. Nonetheless, PhD training, which takes much longer and is more expensive, remains essential for engendering more advanced bioinformatics and data science independent research leadership [33], and was thus also supported on some of the training programs, albeit on a smaller scale (fewer numbers).

Post-graduation pathways

These graduate programs have been successfully started and have taken root in East and West African Universities. Their establishment has included substantial investments by way of careful planning, funding, curriculum development, recruiting, retaining and growing a cohort of instructors, careful recruitment of trainees and extensive efforts in instruction and mentorship [34–36]. The information covering those programs' development, nature and delivery have been well documented elsewhere [37–40]. However, as with all education and scientific initiatives, the ultimate value has to be measured by first the rates of matriculation and graduation from the programs, second, the post-graduation pathways and placements of trainees and ultimately, their impact on the biomedical research and health landscape. Those post-graduation parameters constitute the focus of this paper.

Materials and Methods

The broad training program was initiated under the Human Heredity and Health in Africa (H3Africa) [41–43] Global Health Bioinformatics Research Training Program. The goal was to build genomics research capacity by supporting bioinformatics research training programs at LMIC institutions within Africa. It emphasized partnership with other African or high-income country (HIC)–based collaborators to build sustainable centers of bioinformatics research training relevant to global health research for the African continent.

Specifically, four training programs were initiated and supported, supporting training at five universities spanning five countries in East and West Africa, with the support of a coordinating center in South Africa (Fig. 1). The support included trainee tuition, stipend, international research placements for PhD students, salary support for instructors, curriculum formulation, training materials, connectivity costs and annual consortium meetings. The four host universities were, in turn, supported by three US universities (Baylor College of Medicine, Georgia Institute of Technology and Tulane University) which provided up to one year of research placements and mentorship for PhD students. The training programs were respectively titled: BRECA based at Makerere University in Uganda; the EANBiT based at the ICIPE in Kenya with training at Pwani University; the WASLITBRe based at CU in Nigeria with additional training at the KNUST, Ghana; and the WACE-B based at the USTTB in Mali. All programs were duly accredited by the local accreditation bodies. Training was supported beyond this core network of host institutions, and there were some shared activities between some programs and countries, especially around curriculum design. The EANBiT training program supported the initiation of a bioinformatics MSc program at Pwani University, supported trainees within the Makerere University program, and offered scholarships to several Ugandan students tenable at Pwani University in Kenya; EANBiT and BRECA held joint internship programs and joint curriculum review meetings [37]. WASLITBRe supported trainees at CU in Nigeria, USTTB in Mali and KNUST in Kumasi, Ghana. Makerere University's BRECA trained some students from Eswatini, formally the Kingdom of Eswatini and Kenya. The principal investigators, co-investigators and trainees from all programs met twice a year at the consortium bi-annual meetings facilitated by the coordinating center to share progress, recruitment and training, and jointly evaluate approaches as well as curriculum updates. Together, these accredited training programs constituted the first large-scale, transnationally coordinated and sustained higher-degree level training in bioinformatics, which inevitably required significant resources.

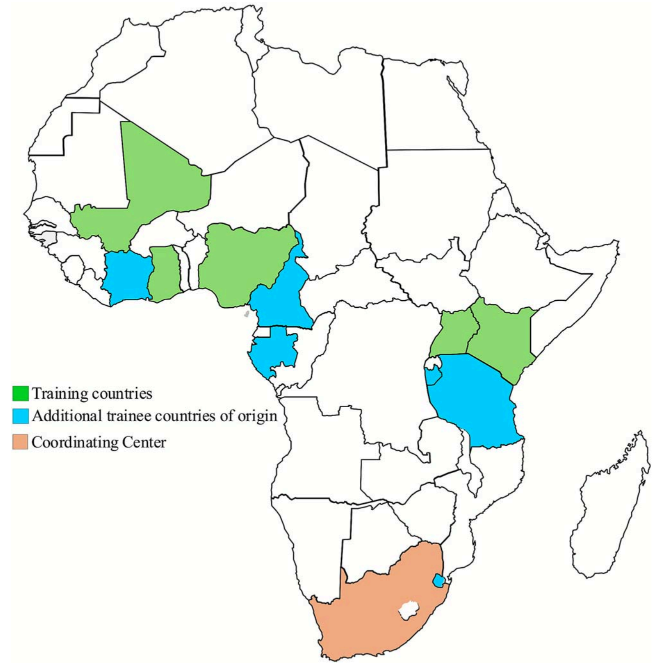

Figure 1 Map of training countries (green), coordinating center (orange), and additional countries of origin for some trainees (blue).

Data gathering

For each of the four training programs, principal investigators and instructors collected four types of data related to the program outputs: (i) they compiled data on the numbers of trainees who had enrolled in both the MSc and/or PhD courses on their programs, (ii) compiled data on numbers of trainees who have graduated from each of those levels, (iii) contacted program graduates and collected data relating to current or previous jobs or placements held by the trainees post-graduation—their roles, institutions and focus of their work by domain, (iv) data on research being performed by various graduates as evidenced by journal publications, some of which were identified using electronic searches like PubMed (grant number or individual name as a key words).

Data analysis

The above-described data were then compiled, curated, collated, categorized, ranked and visualized using R (v4.0.5) programming software to appropriately illustrate the various dimensions of outputs and impact from the training programs. For enrollment and graduation, the numbers of enrolled trainees from each of the five training sites/universities were compiled for each year the program has been running—for both the MSc and PhD trainees. For graduation, a similar compilation was then done for each of the five training sites/universities, starting 2–3 years after matriculation, when the trainees had started graduating (Table 1).

Tracking post-graduation pathways and placements

To establish the postgraduation pathways, the principal investigators and instructors of all four programs respectively conducted a survey that reached out to not only program graduates but also to all persons who had initially enrolled in the programs. Information was collected

Table 1 Number of trainees enrolled and graduated for each training site at the different levels (MSc or PhD) for the different years. Trainees that matriculated from 2018-2021 for MSc and 2018-2020 for PhD are due for graduation and those that matriculated 2022-2024 for MSc and 2021-2024 for PhD were still in training at the time of this assessment.

MSc

Enrolled									**Graduated to date**				
Training sites	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**	**Total**	**2021**	**2022**	**2023**	**2024**	**Total**
BRECA-Uganda		28	16	12	13	17	20	**106**		19	7	12	38
EANBiT-Kenya	8	12	12	12	0	11	7	**62**	8	7	10	17	42
WACE-B-Mali		3	5	10	10	2		**30**	10	10	2		22
WASLITBRe-Nigeria		1		2				**3**			1		1
WASLITBRe-Ghana	0	4	0	5	0	10		**19**		4		5	9
Total/Year	**8**	**48**	**33**	**41**	**23**	**40**	**27**	**220**	**18**	**40**	**20**	**34**	**112**

PhD

Admitted									**Graduated to date**				
Training sites	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**	**Total**	**2021**	**2022**	**2023**	**2024**	**Total**
BRECA-Uganda		6	1	3	3	4	4	**21**			2	2	4
EANBiT-Kenya													
WASLITBRe-Nigeria	8		1	2				**11**	1	2	4	2	9
WACE-B-Mali			5	5	5			**15**				4	4
WASLITBRe-Ghana	0	4	0	1	0	1		**6**				3	3
Total/Year	**8**	**10**	**7**	**11**	**8**	**5**	**4**	**53**	**1**	**2**	**6**	**11**	**20**

about where (institutions or industries) these graduates were now working and brief descriptions of research projects or industry programs they were working on.

Measuring research impact of graduates

Beyond establishing the post-graduate placements of the trainees, we also attempted to measure the extent of their impact, particularly in research informing biomedical and public health progress. This was measured by peer-reviewed publications they had authored or co-authored prior to or after graduation. Information about research domains impacted by their work, their institutional affiliations, diseases investigated and extent of cross-country collaborations was then abstracted from the publications.

Results

Trends of enrollment

The five training sites/universities had different rates of admissions and graduations over the 7 years under assessment. All programs used at least the first year of the grant period (2017–2018) to develop and pursue accreditation of the curricula, recruit instructors and advertise the programs. Consequently, actual instruction started after the 1st or 2nd year, depending on how fast the accreditation process was at the different universities. Hence, graduation did not start until the 4th year of the program. Over the seven years, the cumulative number of trainees admitted to all programs totaled 220 for the master's and 53 for the PhD. The close to 4:1 ratio of master's versus PhD fellows is to be expected, given that the doctoral training is more expensive, takes significantly longer, and has the completion of a master's as one of the requirements for admission, resulting in a substantially smaller pool of qualifying applicants. On average, ∼35% of admitted trainees were female, about one in every three trainees. This is a comparatively remarkable percentage, given that in Africa, the rate of females in science, technology, engineering, and mathematics (STEM) graduate training is significantly lower [44].

Trends of graduation

Ninety (90) of the 220 trainees admitted to the masters program joined in the last three years and are thus still under training. However, of the 130 trainees admitted for the master's program over the first four years (and should therefore have graduated), a total of 101 have graduated across the five training sites. This represents a graduation rate of ∼77%. Several factors account for the remaining 23% that have not graduated. Some, especially among the self-supporting trainees, dropped out of the training due to financial constraints (9%). Others are taking longer to complete their research projects due to delayed experiments or data acquisition and are therefore still in the training pipeline (5%). A few, for various personal and family reasons, deferred their studies (9%). Notably, the graduation rates among those with fellowships are significantly higher at ∼90%. This sub-group not only has fewer financial roadblocks but is also pre-selected by way of taking applicants with the highest potential for competitive fellowships. Regardless, the 77% overall completion rate, particularly in STEM, compares favorably to graduation rates elsewhere [45], such as 74%–77% in STEM fields at U.S. public research universities [46]. Of the 25 trainees admitted for PhD over the first three years of all the programs and should therefore have graduated, 17 have graduated. This represents a graduation rate of 68%, which compares well with graduation rates elsewhere. This is likely because all PhD trainees had fellowships of some sort and were thus unhindered by financial constraints. Also, doctoral fellows tend to be more committed and focused, a criterion assessed during selection for the scholarships through personal statements and oral interviews. The remaining 32% is almost entirely accounted for because their research projects are taking longer due to experimental validations, data acquisition and data analysis delays. For a few, it is taking longer to get their publications through the peer review process—at least

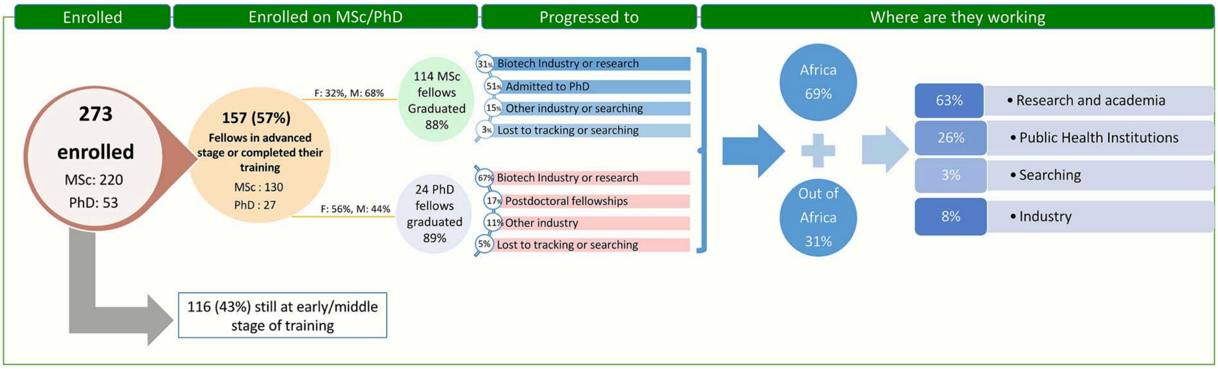

Figure 2 Pathways and placements of program trainees, from enrollment to current positions and work.

one first-author publication is a prerequisite for PhD graduation in the programs.

Post-graduation pathways and placements

The importance of training cannot be overstated, but it is only the first step. The value chain is only completed when the trained students transition into meaningful, gainful and impactful employment and placements. Therefore, having had some cohorts of graduations from all the programs, we sought to assess their post-graduation pathways, placements and career trajectories. Data relating to where graduates of these programs end up, the type of work they are doing there and the impact they are having was collected and analyzed as described in the methods section. The type and trajectories of their post-graduate pathways were found to be both diverse and dynamic. Diverse by way of the nature of work they do in terms of methods and approaches they deploy, the specific diseases and/or public health problems they are working on, the type of institutions and establishments they serve, and their geographic destinations. Dynamic because even a few years post-graduation, some have already changed jobs or areas of focus. For MSc trainees, such placements tended to take the shape of placements on PhD programs, positions as research officers at biomedical and other research institutes, disease surveillance at Ministries of Health and other public health establishments, positions as research officers on funded research projects, and in some cases, successfully applying for, and being awarded independent small grants. For PhD graduates, the pathways tended to majorly lead to postdoc positions. However, some obtained academic positions, independent investigatorships and lead roles on existing or new research grants at universities or research institutes.

About 50% of PhD graduates have ended up in competitive postdoc positions while simultaneously serving as instructors and mentors of master's students. Indeed, almost all PhD students and graduates are currently instructors or mentors of the masters programs. This should, in due course, facilitate their transition to academic and/or independent research positions, which will additionally support the sustainability of the programs by accruing new instructors and mentors. The high percentage of PhD graduates taking the postdoc route is not surprising since postdoc positions tend to be a common next step for all doctoral graduates. What is remarkable is that all have been in competitive positions, several at top international institutions, which at least partly speaks to the quality of training they received. One-quarter of PhD graduates have ended up at institutions outside Africa, primarily in Western Europe and North America. This could be seen as brain drain given the dire need for their skills on the continent; it also represents an opportunity for future trans-continental collaborations and a potential channel for skills and knowledge transfer if they eventually return to the continent or continue to collaborate with colleagues in Africa as some have done. It also speaks to the increasingly global nature of the scientific enterprise and Africa's involvement through its scientific diaspora. Approximately 50% of PhD graduates have taken non-postdoc positions at research institutes, public health laboratories or industry (Fig. 2). Approximately 51% of MSc graduates have ended up in competitive doctoral positions within these same training programs or at universities elsewhere, both within and outside Africa. That half of MSc graduates choose the PhD track could be seen as a proxy for the success of the programs in having trainees appreciate the specialty enough to want to commit to it as a career. It is also a boon to the long-term goal of strengthening scientific leadership on the continent via independent research. The remaining MSc graduates have taken on research positions at biomedical research institutes, public health laboratories, and other industries, including non-science tracks. Remarkably, a few have won competitive small research grants from international organizations such as RSTMH or received competitive research positions on international consortia such as the Pan-African Mosquito Control Association (PAMCA), IMMERSE-U and H3ABioNet.

For graduates that have ended up at public health laboratories or ministries of health, the impact has been more direct in terms of enhanced national disease surveillance programs and improved public testing, diagnosis and data analysis. At the Uganda National Health Laboratory Service (UNHLS), for example, a graduate of the program analyses high throughput sequence data for HIV, SARS-CoV2, *Plasmodium falciparum*, cancer, and Gram-negative enterobacteria. Another performs genomics-level analysis for the Ebola virus, Monkeypox, *Klebsiella pneumoniae*, and *Vibrio cholerae*. The two have additionally developed functional bioinformatics tools, including the published HIV-DRIVES program [47]. Yet another supported building a pipeline for detecting minority HIV resistance variants [48]. A similar pattern was observed at the Uganda Virus Research Institute (UVRI), where program graduates, with the mentorship of their instructors, have been involved in HIV research [48–51] and building bioinformatics HIV infrastructure like customized platforms for HIV sequence storage and analysis for drug resistance [48, 49]. At the Kenya Medical Research Institute (KEMRI), a graduate is researching the population structure pre- and post-vaccine after pneumococcal vaccination, while another is a Research Associate at the International Livestock Research Institute (ILRI), and yet another is a bioinformatics engineer at GeneNetwork at the University of Tennessee Health Science Center. In Nigeria, a graduate works as an infectious disease modeler at Corona Management Systems and an adjunct researcher at CU. These institutions were previously hard-pressed to do this type of work, and typifies the tracks these graduates are taking. The complete list of placements and current work of graduated trainees can be seen in Appendix A.

Institutions impacted or served by graduates of the programs have ranged from research institutes, public health laboratories and universities to hospitals and ministries of health. Of note is that some graduates have worked across a multiplicity of institutions simultaneously or serially, partly because their skills are still scarce in the ecosystem. Thus, several new linkages and institutional partnerships have been enabled or strengthened (Fig. 3a and b).

Beyond the impact the program graduates are having at their respective destinations, the programs themselves have left substantial capacity at their various universities in their wake. MSc and PhD programs in Bioinformatics have been accredited and established. A critical mass of bioinformatics instructors is building up as graduates from the programs pick up instructorships and lectureships on the program or instructors from elsewhere are attracted either as full-time or adjunct staff. Currently, staffing levels are: 18 at Makerere University, 9 at Pwani University, 10 at USTTB, 7 at CU, and 12 at KNUST. Significantly, the programs were institutionalized and have continued to be self-sustaining after the expiry of the grant periods, either through fellowships to trainees from other funders or by direct tuition payments from self-sponsoring students. In Makerere University for example, 65% of current enrollees are self-sponsored, while 35% are sponsored by independent grants or institutions. Unsurprisingly, the programs have proved a boon to the instructors themselves, providing a platform for enhancing their pedagogical and research skills, a source for junior colleagues to support their research projects, and a network to establish new collaborations and find research grants. Between them, instructors in these programs have won awards for >24 research projects or training programs (Appendix B), and several have attained promotions or new appointments, at least in part due to their work, network and support from the programs. This improvement in instructor profiles is an essential dimension of the programs' influence on local research capacity building.

Research contributions

The research outputs and consequent research impact by program graduates and instructors are numerous and diverse, especially as measured by the peer-reviewed research publications they have authored or co-authored. They currently exceed 110 publications, covering four broad domain (Table 2).

About 71 of those papers are about diseases of importance to Africa, 28 papers relate to bioinformatics methods (software, pipelines and algorithms), 13 are reports on bioinformatics programs—training approaches and frameworks, and six relate to the ethics and policy considerations of the field's application in health. Of note, these research outputs were done in largely collaborative contexts, where program participants were first authors on a majority of the papers (53%), second authors on 37% and last authors on 37% (Table 2). This represents a substantial contribution to that work and fairly equal collaborative relationships (Fig. 4).

Research on diseases of importance to Africa

As might be expected, and perhaps appropriately so, a considerably higher volume of the graduates' research has leaned toward diseases of importance to the continent. It ranges from those responsible for the highest morbidity and mortality (Malaria, HIV and TB) to neglected tropical diseases like leishmaniasis. In between are emerging and re-emerging diseases and pandemics like COVID-19, viral hemorrhagic fevers like arboviruses and rhinoviruses; a range of NCDs of growing prevalence like diabetes and high blood pressure (Hbp), as well as NCDs that tend to have a unique prognosis and outcomes among Africans such as cancer, preeclampsia and kidney disease (Fig. 5). Covered on a lower scale are neural and mental disorders like epilepsy, autism, post-traumatic stress disorders and meningitis (Fig. 5). An attempt has been made by program alumni to evaluate the extent of genomics resources in Africa, providing a catalog of the current landscape of sequenced and publicly shared pathogens in different countries on the continent [52].

Malaria

For malaria, research findings have ranged from the establishment of a treatment protocol for the prevention of uncomplicated *P. falciparum* malaria in the demographic where most mortality lies—children between 3 months and 9 years old [53]; a description of the contribution of travel to malaria transmission [54]; the identification of factors associated with the occurrence of clinical malaria across different ecological settings in Mali [55] which could help the development of new strategies for malaria elimination; and the establishment of a multiplicity of infections in asymptomatic malaria patients [56]. Several studies have been performed in the malaria therapeutics space for potential drug targets [57–64], including the use of pharmacophore modeling to identify important potential lead compounds that could serve as inhibitors of Pf 5-ALAS [56] and the use of molecular docking techniques to identify potential inhibitors of PfPTPS [62]. Other studies have explored Anopheles and Plasmodium metabolic products, pathways and gene usage as well as transmission, development and function of the vector and parasite [65–69].

Tuberculosis

TB research has included a review examining the current progress in studying polymorphisms within immune genes associated with TB susceptibility, focusing on African populations [70]. It has also investigated the impact of *Mycobacterium tuberculosis (M.tb)* infection and *M. tb*-specific IFNg immune responses on airway microbiome diversity. This was done by performing TB GeneXpert and QuantiFERON-GOLD assays during the follow-up phase of a longitudinal HIV-Lung Microbiome cohort of individuals recruited from two large independent cohorts in rural Uganda [71]. Closely related work found that airway microbiome signature accurately discriminates *M.tb* infection status [71, 72]. Further work is now leveraging machine learning to predict antibiotic resistance in four anti-TB drugs (rifampicin, isoniazid, streptomycin, and ethambutol) as part of disease surveillance [73].

Human immunodeficiency virus

Work in HIV has been varied, ranging from a mapping of the distribution of HIV subtypes in several African countries [74] to the molecular epidemiology and phylogenetics-based studies of transmission, including the finding that men disproportionately contribute to HIV

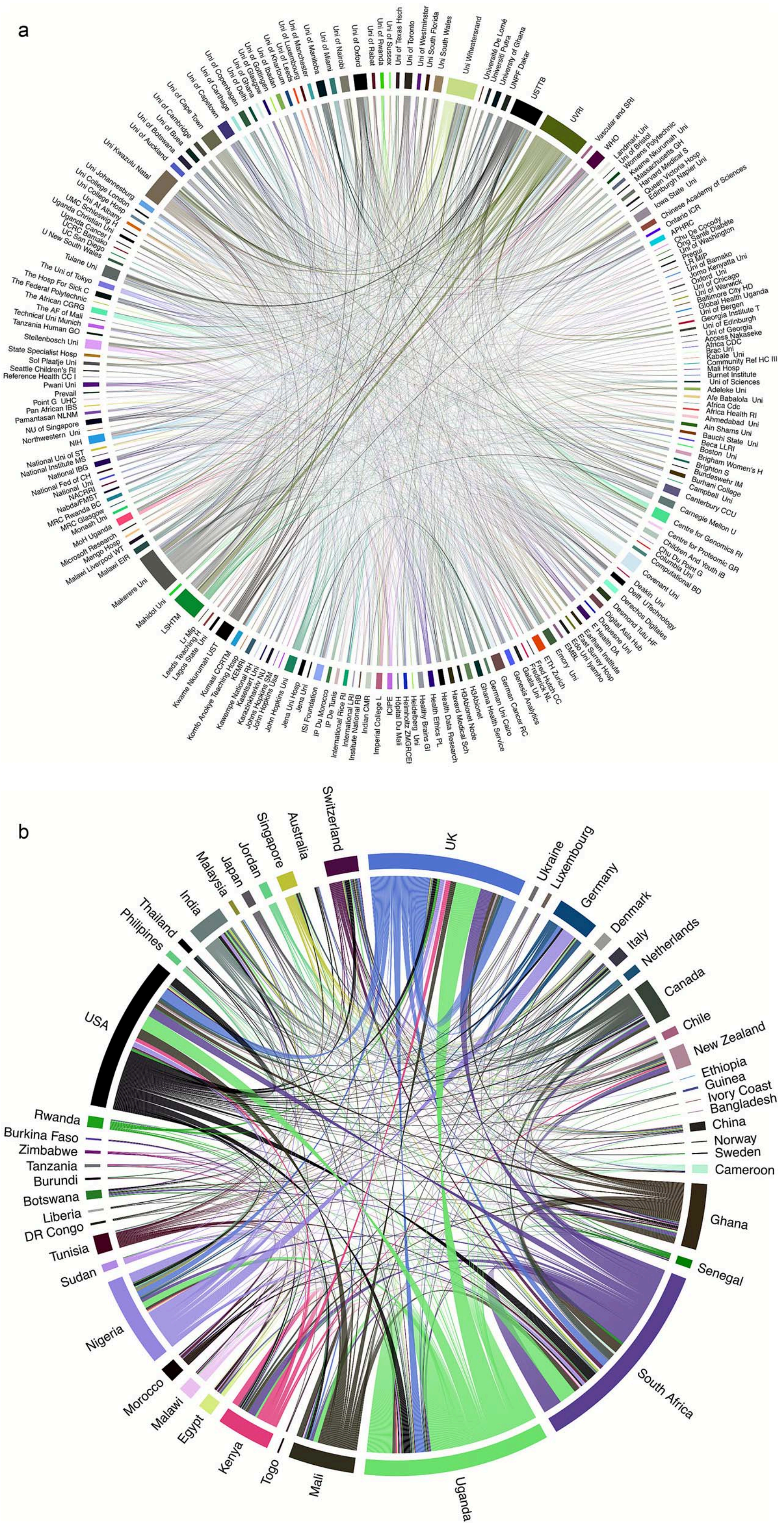

Figure 3 (a) Inter-institutional collaborations as derived from publications for which program participants were authors or co-authors. (b) Transnational collaborations as derived from collaborative publications.

transmission and should thus be mainly targeted in its management from a public health perspective [75, 76]. Other transmission-related work showed that pregnant women have cervicotype 4 (CT4) vaginal microbiota, which is associated with inflammation, itself a known catalyst for HIV acquisition and transmission [77]. It has also involved the development of custom tools for HIV sequence analysis, including

Table 2 Broad categories of domains covered by graduates' work.

Domain	Publications	Percentage
Diseases	71	60.2
Bioinformatics methods and tools	28	23.7
Bioinformatics capacity building programs	13	11.0
Ethics and policy	6	5.1

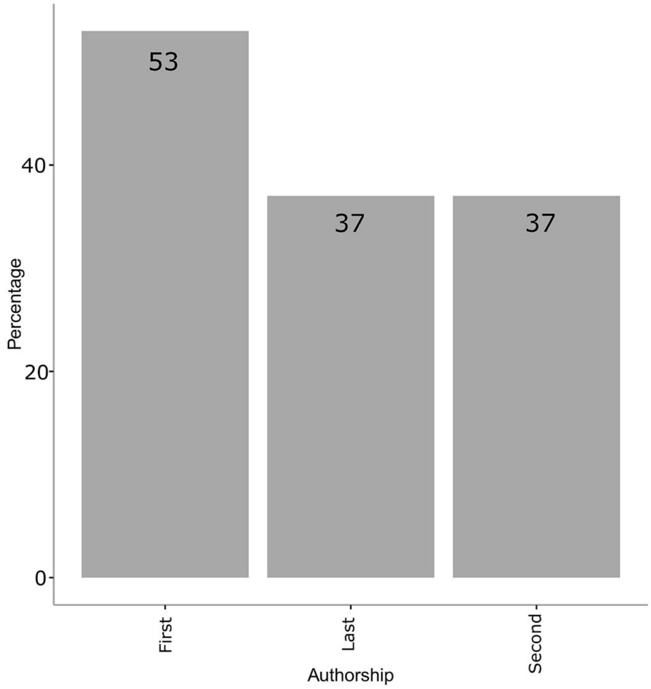

Figure 4 Percentage of first, second, and last authorships contributed by program participants on collaborative publications ($N = 110$). Participants most frequently appeared as first authors (53%), followed by second (37%) and last authors (37%), indicating strong involvement in both lead and senior research roles.

developing a pipeline that detects mutations in HIV RNA-Seq data [48] and another that performs HIV drug resistance identification, variant evaluation, and surveillance [47]. Another tool integrates NGS, clinical and demographic data for analysis of HIV drug resistance [49]. Further methods-based work applied fine-scale network analysis to estimate transmission network parameters of HIV sequences from key populations, revealing the potential value of such analysis in predicting the progression of the HIV epidemic [50]. Additional research has explored the use of metagenomics as a tool to study non-malarial febrile fevers in HIV-infected children [78].

Emerging and re-emerging infections

Research by program fellows has fittingly investigated some emerging and re-emerging diseases and pandemics on the continent. Even though Africa had comparatively fewer COVID-19 fatalities [79], the pandemic received some research attention from program fellows. For example, there was an assessment of the dynamics of circulating SARS-CoV-2 variants identified during the different COVID-19 waves in Mali [80]. Similarly, the need for continuous surveillance of SARS-CoV-2 virus to detect emerging variants was illustrated in Uganda [81]. Relatedly, establishing the demographics and seroprevalence of the virus in vaccine-naïve participants was performed in the Democratic Republic of Congo (DRC), Guinea, Liberia, and Mali [82]. There have also been attempts to research potential therapeutic targets for the disease, including an application of immuno-informatics that identified four proteins with advantageous binding to the relevant human HLA-1 and a use of virtual screening of African natural products databases to identify high-affinity inhibitors of SARS-CoV-2 main protease [83]. Beyond the biology of COVID-19 itself, the lockdowns during the pandemic provided an opportunity to confirm that air quality can be used as a proxy for population mobility [84] and to determine the effectiveness of virtual reality (VR) as a tool for training frontline health workers in the context of highly infectious diseases like COVID-19 [85].

Work on other emerging and re-emerging diseases has highlighted the need for systematic molecular studies to understand the transmission of Q fever [86] and the development of a type-specific full genome sequencing approach for obtaining Human Rhino Virus A (HRV) genomic data to characterize infections [87]. For arbovirus infections, a program graduate has contributed to identifying early infection molecular signatures that can differentiate patients who will progress to severe dengue infection. These findings could pave way for evaluating of immunomodulatory therapies for dengue [88] (under review). Additional work has described mosquito-borne viruses in Uganda, focusing on their discovery, morbidity, mode of transmission, animal hosts and clinical manifestations when involved in disease outbreaks. Work on other common viral infections has covered developing and validating a method for detecting and quantifying hepatitis b virus (HBV) DNA by qPCR [89] and, importantly, using NGS to highlight the presence of undiagnosed viruses causing measles-like illnesses (MLI) in Uganda, including vaccine-preventable illnesses. Such findings inform diagnostic assay selection and vaccination strategies [90].

Non-communicable diseases

A range of NCDs of growing prevalence like diabetes and Hbp, as well as NCDs that tend to have a generally distinct prognosis and outcomes among Africans such as cancer, preeclampsia, sickle cell disease and kidney disease have been investigated. For cardio-vascular diseases for example, multivariate meta-analysis has helped identify novel genetic loci associated with cardiometabolic risk factors (CMRFs) in two continental African populations—Ugandans and South Africans [91] and some genetic variants associated with Hbp traits have been identified [92]. Surveys have established the prevalence of type 2 diabetes (T2DM), community awareness of the condition and its attendant risk factors in rural Mali [93], as well as the development of a set of materials for its management [94]. Molecular dynamics simulations have been used to gain insights into the conformational stability of the insulin receptor gene to better understand the physiology of the disease [95]. Its intersection with kidney disease has been investigated through a determination of whether inhibition of kidney function drug targets adversely impacts T2DM [96] as indeed has its intersection with dyslipidemia [97]. Genetic variants associated with lipid traits in African populations have been identified [98]. For kidney disease, which tends to have a unique prognosis among African ancestry populations, binding sites within the APOL1 protein that could be an attractive site for potential inhibitors of APOL1 in the

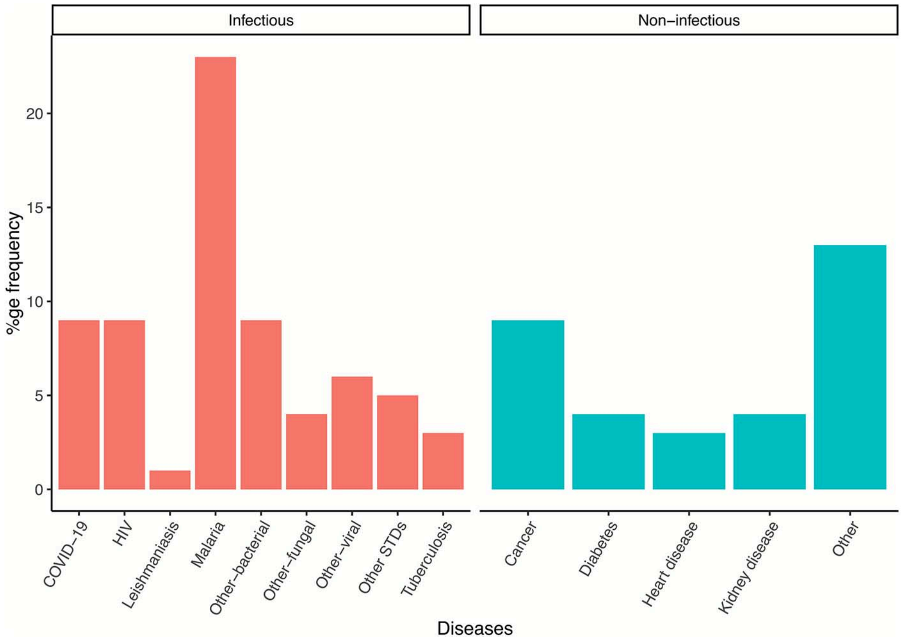

Figure 5 Relative frequency of research on different diseases of public importance on the continent.

management of chronic kidney disease were located [99] as were TrkB agonists for the treatment of CDKL5-deficiency disorders [100]. Uganda's largest patient consented registry was established for sickle cell disease, enrolling patients across four regional referral hospitals. This registry supports studies on clinical phenotypes, outcomes, and interventions like hydroxyurea use and newborn screening showcasing the feasibility of longitudinal disease registries in resource-limited settings [101].

Cancer

As a growing epidemic on the continent [102–104], cancer has received a decent amount of attention from program graduates. That ranges from an assessment showing low knowledge and perceptions of urinary blood cancer among dye workers in Ghana [105] to a review of studies that used radiation therapy as a treatment for breast cancer in Africa, focusing on survival outcomes, adverse effects, radiation therapy techniques, fractionation schedules, and effectiveness of radiation therapy [106]. Cancer studies have similarly explored the different clusters of microbial communities that form different functional groups in gastric cancer [107] and antimicrobial drug resistance, exemplified by the genotypic and phenotypic resistance to third-generation cephalosporin in hematologic cancer patients [108]. Other studies have identified potential biomarkers that can be used for the prognosis of Burkitt's lymphoma [109], first described in Uganda by Dennis Burkitt as a lymphoid malignancy involving mostly the jaw of children from equatorial Africa [110]. Further cancer-related work identified essential genes that could be exploited as potential biomarkers for prognosis and therapeutic targets in hematologic malignancies [111].

Maternal health and mental disorders

The training has additionally supported some nascent work on maternal health, for example the identification of three candidate biomarkers, sFlt1, sFlt1/PlGF ratio and PIGF, for the diagnosis of preeclampsia in pregnant women [112] and determination of the prevalence of osteoporosis in African postmenopausal women [113]. Work on mental disorders has covered a review of functional studies and the clinical impact of the pathogenic mutations in the sodium channel genes that contribute to epilepsy [114]. Additional research explored the association of plasma protein levels with major depressive disorder (MDD) in multi-ancestry studies across diverse populations [115]. Similarly, analytical work has been applied to identify genes that significantly contribute to autism and a corresponding classifier has been developed to predict the severity of the disease in the neonate [116]. Further work on mental disorders has explored the interaction of post-traumatic stress and ischemic stroke, showing differences in the genetic susceptibility to the development of post-traumatic stress and its associated risks for ischemic stroke between individuals of African and European ancestries [117].

Antimicrobial drug resistance and therapeutics

As one of the most existential problems facing public health [118], some work has been devoted to AMR and therapeutics. This has included the development of arylidene-based quinazolin-4(3H)-one motifs as potential antimicrobial drug candidates for a range of bacterial and fungal pathogens—*Staphylococcus aureus*, *Candida albicans*, *Aspergillus niger* and *Rhizopus nigricans* [119]. This is important work since *C. albicans* and *Staphylococcus* species are the most common fungal and bacterial agents isolated from bloodstream infections worldwide [120]. For *S. aureus*, further work has involved an evaluation of its antibiotic resistance patterns and virulence factors in isolates from pleuritis infections of hospitalized patients [121]. Disturbingly, AMR in gonorrhea has been characterized in Ugandan men, alongside poor antibiotic stewardship and near-universal resistance of key antigonococcal antibiotics [122]. Appropriately, a standardized gonococcal antimicrobial resistance surveillance program has been established in the country to help support population-level interventions [123, 124] including the development of a portable magnetofluidic platform for effective testing of sexually transmitted diseases like gonorrhea and their susceptibility to drugs [125]. More recent work has explored bringing Machine Learning and data-driven approaches to bear on AMR surveillance. Specifically, this has included testing/calibrating the effectiveness of models trained on more readily available European data for predicting AMR in Africa [12] and optimizing antibiotic use to combat antimicrobial-resistant infections in Uganda and other LMICS through Centers for Antimicrobial Optimization Network (CAMO-Net) [126]. There have been various efforts to identify potential drugs or drug targets on the therapeutic front. For example, *in silico* methods like molecular docking have been applied to investigate alternate antibacterial agents for Methicillin-resistant *S. aureus* from a library of benzimidazole derivatives [127] and a computer-aided drug design approach to predict potential inhibitors of *Salmonella typhimurium*—ortho acetyl sulphydrylase synthase (StOASS) [128]. Simultaneously, microwave technology has been deployed to develop coumarin-based compounds that could be used in drug development [129].

Bioinformatics methods, software, pipelines, and algorithms

Work by program graduates has stretched beyond the etiology, pathology and treatment of diseases to include the development of computational methods, approaches, frameworks and techniques for analyzing various biomedical and public health data. For example, a computational model for the prediction of new drug targets in Plasmodium was developed [130], and machine learning has been variously applied to develop predictive models for identifying 208 dependency factors (HDFs) in Drosophila involved in disease pathogenesis [131], predicting essential genes in Drosophila [132] and identifying several essential genes across eukaryotes [133] and predicting foot-and-mouth disease outbreaks in cattle [134]. Indeed, a full review of the standard procedures and resources available for predicting essential genes in organisms using machine learning has been done [135]. Further methods work has involved the development of a system of ordinary differential equations to model malaria transmission dynamics [136] and a review of clustering methods for big data [137]. Some work was committed to evaluating imputation panels and tools for HLA genes [138], yet other work reviewed methods for studying host–parasite protein–protein interactions in malaria [139]. That review provides a useful knowledge base. In terms of experimental methods, fellows' work revealed novel insights into the potential use of black soldier fly larvae (BSFL) in the bioconversion of various highly lignocellulosic diets to fermentable sugars for subsequent value-added products such as bioethanol [140]. It has also included determining the value of an elastic tissue stain in assessing vascular invasion in Colorectal Cancer (CRC) [141]. Finally, a stable and highly accurate numerical tool for computing river flows in urban areas was built as a first step toward a tool for flood predictions [142].

Bioinformatics training approaches and frameworks

As part of self-introspection, program research explored the organization, content and delivery of bioinformatics and data science training at different levels. Important work provided guidelines on how to apply competency-based bioinformatics education and training and illustrated how the competencies apply to different knowledge areas for curriculum development [35, 143]. Another effort designed and implemented a model for bioinformatics training and mentorships in resource-limited settings [39], while another described a mentorship program for improving bioinformatics in Kenya [38]. A similar program involved the Nigerian bioinformatics and genomics network and its impact and importance within and outside that country [144]. Those training frameworks and curricula were used to improve bioinformatics training across the continent, for example within EANBiT, which evolved to meet challenges for experiential learning through mini projects to enhance the acquisition of skills and collaborations [37]. Trainees have also had opportunities to network and connect with others on the continent at continent-wide events such as the 3rd ISCB Africa Student Council Symposium [145]. Besides mentorship programs, hackathons have also been used as tools to teach bioinformatics while solving an actual bioinformatics problem simultaneously [146]. For example, one of the hackathons built two data portals for microbiome and genomics in African populations [146].

Furthermore, having appreciated the need for advanced training in data science and bioinformatics on the continent, there have been efforts to improve the available bioinformatics and data science infrastructure, mainly through public–private partnerships as in the case of the African Centers of Excellence in Bioinformatics and Data-Intensive Sciences (ACE) [147, 148], based in Uganda and Mali, which were integral parts of the training programs. This ACE project aims to enhance computational infrastructure, develop advanced bioinformatics and data science skills among local researchers and students, and provide innovative emerging technologies for infectious diseases research [147]. One such innovative technology has been VR, which has been used to upgrade postgraduate surgical training in resource-limited settings [149] and to train frontline health workers, especially in the context of highly infectious diseases like COVID-19 [85]. The feasibility of other emerging digital solutions including AR, in meeting the challenges of delivering rapid and equitable access to emergency care training at scale has been explored. This has produced specific policy and practice recommendations for health professional education and training to enhance emergency care [150].

Ethics and policy considerations

Research outputs have also covered and considered the important aspect of ethics and policies that accompany the application of such tools as bioinformatics, data science and artificial intelligence to health in general and biomedical research in particular, especially in resource-limited settings. There have been discussions of innovations in the ethical governance of AI, including highlighting areas in urgent need of attention internationally [151]. There has been exploration and advice for data standards, practices and philosophies of AI with their specific interplays in the role of shaping the future of society to ensure that AI systems contribute to a better future for all [152]. Notably, individuals from the training programs have participated in formulating WHO guidelines for use of AI in health [153]. They have further participated in international efforts exploring the ethics of using digital technologies for health promotion, particularly among the youth [154]. In the policy space, there has been work showing the need for African governments to invest in public plant genomics research and applications, supporting bioinformatics platforms and training programs, and stimulating university and industry partnerships in agriculture and medicine [155]. Separate work has shown the need for governments to invest more in RNAseq-based research in Africa [156]. Equally, complementary efforts have described the current state, barriers and opportunities for translational research and showed the need to integrate it into the healthcare system [157].

Discussion

This work assesses the pathways, outputs and impact of NIH-supported degree-level graduate bioinformatics training programs on the African continent—the first ones of their scale in numbers and geographical spread. It does not attempt to describe the structure, implementation, execution and delivery of those programs since those have been variously documented elsewhere in papers that describe individual programs [37–40]. Here, we assess their net outputs and impact in aggregate, covering all four funded projects established at five universities in five countries. Together, they have enrolled dozens of trainees, several of whom have already graduated, several are still in the training pipeline and many more continue to matriculate into the training programs. The graduated trainees have since spread out into multiple career paths and several institutions, where they have made multiple impactful contributions, especially through research, teaching and science leadership. Those contributions have resulted into various publications of importance to African biomedical research and public health, as well as new research and training programs. It is noteworthy that most of that research and resultant publications have been in collaboration with scientists at other institutions on the continent, and those from other continents, particularly North America and Europe. Strengthening collaborative science has thus been one of the major incidental outputs of these training programs, as has been the formation of a viable community of bioinformatics and genomics professionals able to specialize in different research endeavors while simultaneously leveraging the synergies of cross-continental research and training.

The amount and extent of outputs are testimony to the value of sustained long-term graduate training. Unsurprisingly, such training requires substantial initial funding, an availability of a threshold of initial trainers, elaborate curriculum designs, reliable accreditation bureaucracies, quality assurance systems and the existence of a reasonable qualifying pool of potential trainees. They also benefit tremendously from sustained support by hosting institutions and the vision and enthusiasm of faculty. Indeed, the NIH funded the initiation of the training programs, i.e., curriculum design, staff support, students' fellowships (tuition and stipend), teaching materials, travel and coordination. This amounted to ∼\$5.2m [158] over a period of five years across the five training sites, representing ∼75% of the total cost of the programs. Supplementary support came from host universities, which provided lecture rooms, utilities, computing platforms like HPCs, and part of the staff effort, all of which cost above the allowable NIH indirect costs. Additional funding to the universities came from self-sponsored students who contributed subsidized tuition fees. These local contributions bode well for the continued sustainability of the programs. Now that the programs are in place, they can be leveraged to support other forms of training. For example, program graduates and instructors have supported other training modes, such as the short-term Introduction to Bioinformatics Training (IBT) organized by H3ABioNet [15, 36] and Africa CDC's sequencing capacity building for public health laboratories across the continent [159]. Those have, in turn, also served as a preparation platform for aspiring candidates for long-term graduate training, which illustrates the cyclical relationships and synergies such programs have with other forms of training as well as with research and public health endeavors. Obviously, for sustained relevance and continued quality outputs, the programs must be continuously updated to co-opt new growth areas for biomedical research such as artificial intelligence, machine learning and other mostly unmet needs like generative AI and image analysis. Indeed, in subsequent funding, the NIH has supported either new programs for those areas or their integration into existing ones under the Data Science for Health Discovery and Innovation in Africa (DS-I Africa) Initiative. Perhaps one key unexpected dividend of these programs has been the high participation/integration of female trainees, who have accounted for ∼33% of participants, above normal levels in STEM graduate training [160]. Indeed, recent global estimates indicate that female participation rates in many African graduate STEM programs typically range between 20%–30% [161, 162], comparable to ∼28%–38% of STEM graduate students in the US [160]. Against this backdrop, our observed female participation exceeds typical regional benchmarks and slightly surpasses participation levels in many high-income countries, suggesting meaningful progress toward gender equity in advanced STEM training.

To the extent that outputs often correspond to inputs, the initial substantial level of investment by the NIH for a sustained period of 5 years unsurprisingly generated a significant impact. That initial spark was essential for the progress of bioinformatics and genomics on the continent. To the best of our knowledge, none of the prior short-term trainings had yielded such a level of outputs and impact. Furthermore, the programs have been fortunate to leverage the parallel growth of requisite computational infrastructure—software and hardware [16, 34, 147, 159]—essential for their progression and sustainability [163]. Whereas outputs and impact have been substantial, there is still room for improvement, particularly by optimizing quality internships and placements, which link trainees to ongoing and functional research and development programs during training.

Conclusion

This program assessment has not been without limitations. Despite our best efforts to track graduates, tracking every last one and their outputs has proven impossible. As a result, not all outputs and pathways have been captured, but the overwhelming majority are represented here. At its best, this paper outlines a snapshot in time. This is so, first because there have only been a few cohorts of program graduates at this point and secondly because they are still in the very early stages of their careers. Their outputs and impact will increase in quantity and quality as a function of both time and experience. Additionally, several cohorts are still in the training pipeline and others will matriculate to the programs in succeeding years. Consequently, a future reassessment can be expected to yield more outputs and even more significant impact. The present work offers a pilot approach and framework that can, in the future, be improved or expanded for such an exercise in training program assessment. However, the many outputs outlined in this paper should be better seen as an indication of what a difference appropriately resourced and planned training programs at the correct scale can make and not a suggestion that the training need has been fully met. Five training sites/universities in five countries remain quite few compared to the need, both present and anticipated, as genomics, bioinformatics and data science increasingly take on a bigger share of research, development and clinical practice [6].

Key Points

- The National Institutes of Health (NIH)-supported Human Heredity and Health in Africa graduate bioinformatics programs, launched in 2017 across five African universities, have trained >270 students, producing >110 MSc and 20+ PhD graduates, significantly boosting Africa's bioinformatics and genomics capacity.
- Graduates have transitioned into impactful roles in academia, research institutes, public health laboratories, and industry, contributing to disease surveillance, diagnostics, and cutting-edge biomedical research across Africa and beyond.
- The programs have generated >110 peer-reviewed publications, with research spanning infectious diseases, non-communicable diseases, bioinformatics tool development, ethics, and policy, addressing critical health priorities for the continent.
- A sustainable pipeline of bioinformatics instructors and mentors has been established, with many graduates now teaching and leading research, ensuring program continuity beyond initial NIH funding.
- These initiatives have fostered strong intra-African and global research collaborations, strengthened institutional capacity, and demonstrated the long-term value of sustained, degree-level training for scientific leadership in Africa.

Acknowledgements

This work was supported in part with federal funds from the National Institutes of Health under four training grants: BRecA (grant number U2RTW010672), EANBiT (grant number U2RTW010677), WASLITBRe (grant number U2RTW010679), and West African Center of Excellence for Global Health Bioinformatics Research Training (WACE-B (grant number U2RTW010673) and from the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health, Department of Health and Human Services BCBB Support (Services Contract HHSN316201300006W/75N93022F00001) to Guidehouse Digital and Clinical Data Management (Services Contract 47QTCA22D008F-140D0424F0200) to Research Data and Communication Technologies. Supplementary support came from H3ABioNet (grant number U41HG006941), Cancer Genomics and Genomic Data Science for East Africa (grant number D43CA260656), and Genome-wide characterization of complex variants and their phenotypic effects in African populations (grant number U01HG013442), She Data Science (SHEDS) program: Empowering Uganda's Women in Health Data Science: Identifying Barriers, Bridging Knowledge and Innovation for Tangible Impact; through a collaborative agreement with the University of California San Francisco (UCSF) (Agreement No. UFRA-460), Gates Foundation AI Grand Challenges (INV-062695), eLwazi Open Data Science Platform and Coordinating Center (grant number U2CEB032224) and funded in part (SJR) by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Authors contributions

Daudi Jjingo (Led the conceptualization of the study, Performed formal analysis, Managed funding acquisition, Led the project supervision, Drafted the original version of the manuscript), Andrew Walakira (Performed formal analysis, Produced the visualizations in the paper, Drafted the original version of the manuscript), Suhaila Hashim (Carried out data curation), Cisse Cheickna (Carried out data curation), Ronald Galiwango (Carried out data curation, Performed formal analysis, Drafted the original version of the manuscript), Caleb Kibe (Carried out data curation), Florence N. Kivunike (Carried out data curation), Gerald Mboowa (Carried out data curation, Managed funding acquisition, Led the project supervision), Fredrick Elishama Kakembo (Performed formal analysis, Carried out the investigation, Produced the visualizations in the paper, Drafted the original version of the manuscript), Babajide Ayodele (Carried out the investigation), Jean-Baka Domelevo Entfellner (Carried out the investigation), Santie de Villiers (Carried out the investigation), Karen Wambui (Carried out the investigation, Oversaw the project administration), Segun Fatumo (Carried out the investigation), Tinashe Chikowore (Carried out the investigation), John Mukisa (Carried out the investigation), Alfred Ssekagiri (Carried out the investigation), Nicholas Bbosa (Carried out the investigation), Julius Mulindwa (Carried out the investigation), Samuel Kyobe (Carried out the investigation), Mike Nsubuga (Carried out the investigation), Grace Kebirungi (Carried out the investigation), Eric Katagirya (Carried out the investigation), Savannah Mwesigwa (Carried out the investigation), Ibra Lujumba (Carried out the investigation), Rogers Kamulegeya (Carried out the investigation), Samuel Kirimunda (Carried out the investigation), Stephen Kanyerezi (Carried out the investigation), Shahiid Kiyaga (Carried out the investigation), Ivan Sserwadda (Carried out the investigation), Davis Kiberu (Carried out the investigation), Bernard S. Bagaya (Carried out the investigation), Julius Okwir (Carried out the investigation), Patricia Nabisubi (Carried out the investigation), Grace Nabakooza (Carried out the investigation), Mugume Twinamatsiko Atwine (Carried out the investigation), Ricard Sserunjogi (Carried out the investigation), Rolanda Julius (Carried out the investigation), Mariam Quiñones (Carried out the investigation), Meghan McCarthy (Carried out the investigation), Phillip Cruz (Carried out the investigation), Karlynn Noble (Oversaw the project administration), Christopher J. Whalen (Provided the resources), Darrell Hurt (Provided the resources), Maria Y. Giovanni (Provided the resources), Michael Tartakovsky (Provided the resources), Deogratius Ssemwanga (Provided the resources), John M. Kitayimbwa (Provided the resources), Steven J. Reynolds (Provided the resources), Christopher C. Whalen (Provided the resources), Andrew Kambugu (Provided the resources), Neil A. Hanchard (Provided the resources), Li Jian (Led the project supervision), Peter Amoako-Yirenkyi (Led the project supervision), Graeme Mardon (Led the project supervision), I. King Jordan (Led the project supervision), Samson Pandam Salifu (Led the project supervision), Mamadou Wele (Led the project supervision), Ezekiel Adebiyi (Led the project supervision), Jeffrey G. Shaffer (Led the project supervision), Seydou Doumbia (Led the project supervision), David Patrick Kateete (Led the project supervision), Michelle Skelton (Led the project supervision), Nicola Mulder (Led the project supervision), Jonathan K. Kayondo (Led the project supervision), Daniel Masiga (Led the project supervision), and H3A (Led the project supervision)

Supplementary material

Supplementary material is available at *Briefings in Bioinformatics* online.

Data availability

All data supporting the findings of this study are provided within the manuscript and its supplementary materials.

References

1. Bryson JM, Bishop-Williams KE, Berrang-Ford L. *et al.* Neglected tropical diseases in the context of climate change in East Africa: a systematic scoping review. *Am J Trop Med Hyg* 2020;**102**:1443–54. https://doi.org/10.4269/ajtmh.19-0380
2. George NS, David SC, Nabiryo M. *et al.* Addressing neglected tropical diseases in Africa: a health equity perspective. *Glob Health Res Policy* 2023;**8**:30. https://doi.org/10.1186/s41256-023-00314-1
3. Buscaglia CA, Kissinger JC, Agüero F. Neglected tropical diseases in the post-genomic era. *Trends Genet* 2015;**31**:539–55. https://doi.org/10.1016/j.tig.2015.06.002
4. Chikowore T, Kamiza AB, Oduaran OH. *et al.* Non-communicable diseases pandemic and precision medicine: is Africa ready? *EBioMedicine* 2021;**65**:103260. https://doi.org/10.1016/j.ebiom.2021.103260
5. Mudie K, Jin MM, Tan. *et al.* Non-communicable diseases in sub-Saharan Africa: a scoping review of large cohort studies. *J Glob Health* 2019;**9**:020409. https://doi.org/10.7189/jogh.09.020409
6. Vamathevan J, Birney E. A review of recent advances in translational bioinformatics: bridges from biology to medicine. *Yearb Med Inform* 2017;**26**:178–87. https://doi.org/10.15265/IY-2017-017
7. Brusic V. The growth of bioinformatics. *Brief Bioinform* 2007;**8**:69–70.
8. Ibrahim B, McMahon D, Hufsky F. *et al.* A new era of virus bioinformatics. *Virus Res* 2018;**251**:86–90. https://doi.org/10.1016/j.virusres.2018.05.009
9. Levy SE, Boone BE. Next-generation sequencing strategies. *Cold Spring Harb Perspect Med* 2019;**9**:a025791. https://doi.org/10.1101/cshperspect.a025791
10. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. *Mol Cell* 2015;**58**:586–97. https://doi.org/10.1016/j.molcel.2015.05.004
11. Elemento O, Leslie C, Lundin J. *et al.* Artificial intelligence in cancer research, diagnosis and therapy. *Nat Rev Cancer* 2021;**21**:747–52. https://doi.org/10.1038/s41568-021-00399-1
12. Nsubuga M, Galiwango R, Jjingo D. *et al.* Generalizability of machine learning in predicting antimicrobial resistance in E. coli: a multi-country case study in Africa. *BMC Genomics* 2024;**25**:287. https://doi.org/10.1186/s12864-024-10214-4
13. Bonham VL, Green ED. The genomics workforce must become more diverse: a strategic imperative. *Am J Hum Genet* 2021;**108**:3–7. https://doi.org/10.1016/j.ajhg.2020.12.013
14. Hamdi Y, Zass L, Othman H. *et al.* Human OMICs and computational biology research in Africa: current challenges and prospects. *Omics J Integr Biol* 2021;**25**:213–33.
15. Mulder NJ, Adebiyi E, Alami R. *et al.* H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. *Genome Res* 2016;**26**:271–7. https://doi.org/10.1101/gr.196295.115
16. Kumuthini J, Zass L, Panji S. *et al.* The H3ABioNet helpdesk: an online bioinformatics resource, enhancing Africa's capacity for genomics research. *BMC Bioinformatics* 2019;**20**:741.
17. Attwood TK, Blackford S, Brazas MD. *et al.* A global perspective on evolving bioinformatics and data science training needs. *Brief Bioinform* 2019;**20**:398–404. https://doi.org/10.1093/bib/bbx100
18. Via A, de Las Rivas J, Attwood TK. *et al.* Ten simple rules for developing a short bioinformatics training course. *PLoS Comput Biol* 2011;**7**. https://doi.org/10.1371/journal.pcbi.1002245
19. Lawlor B, Walsh P. Engineering bioinformatics: building reliability, performance and productivity into bioinformatics software. *Bioengineered* 2015;**6**:193–203. https://doi.org/10.1080/21655979.2015.1050162
20. Schneider MV, Watson J, Attwood T. *et al.* Bioinformatics training: a review of challenges, actions and support requirements. *Brief Bioinform* 2010;**11**:544–51. https://doi.org/10.1093/bib/bbq021
21. Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? A proposed definition and overview of the field. *Methods Inf Med* 2001;**40**:346–58.
22. Wood L, Gebhardt P. Bioinformatics goes to school—new avenues for teaching contemporary biology. *PLoS Comput Biol* 2013;**9**. https://doi.org/10.1371/journal.pcbi.1003089
23. Gallagher SR, Coon W, Donley K. *et al.* A first attempt to bring computational biology into advanced high school biology classrooms. *PLoS Comput Biol* 2011;**7**. https://doi.org/10.1371/journal.pcbi.1002244
24. Machluf Y, Yarden A. Integrating bioinformatics into senior high school: design principles and implications. *Brief Bioinform* 2013;**14**:648–60. https://doi.org/10.1093/bib/bbt030
25. Wefer SH, Sheppard K. Bioinformatics in high school biology curricula: a study of state science standards. *CBE—Life Sci Educ* 2008;**7**:155–62.
26. Madlung A. Assessing an effective undergraduate module teaching applied bioinformatics to biology students. *PLoS Comput Biol* 2018;**14**. https://doi.org/10.1371/journal.pcbi.1005872
27. Sayres MAW, Hauser C, Sierk M. *et al.* Bioinformatics core competencies for undergraduate life sciences education. *PloS One* 2018;**13**. https://doi.org/10.1371/journal.pone.0196878
28. Machanick P, Tastan Bishop Ö. How to establish a bioinformatics postgraduate degree programme—a case study from South Africa. *Brief Bioinform* 2015;**16**:346–54. https://doi.org/10.1093/bib/bbu014
29. Işık EB, Brazas MD, Schwartz R. *et al.* Grand challenges in bioinformatics education and training. *Nat Biotechnol* 2023;**41**:1171–4. https://doi.org/10.1038/s41587-023-01891-9
30. Nelson MI, Lloyd-Smith JO, Simonsen L. *et al.* Fogarty International Center collaborative networks in infectious disease modeling: lessons learnt in research and capacity building. *Epidemics* 2019;**26**:116–27. https://doi.org/10.1016/j.epidem.2018.10.004
31. Porter JE. The Fogarty International Center is essential to global health security. *Am J Trop Med Hyg* 2017;**97**:627–8. https://doi.org/10.4269/ajtmh.17-0597
32. Mwangi EM, Iseme-Ondiek RA, Riang'a RM. *et al.* Addressing barriers to post-graduate research training in low resource settings: an innovative approach in an institution of higher learning in Kenya. *Front Med* 2024;**11**:1470922. https://doi.org/10.3389/fmed.2024.1470922
33. Gardner SK. 'What's too much and what's too little?': the process of becoming an independent researcher in doctoral education. *J High Educ* 2008;**79**:326–50.
34. Mulder NJ, Adebiyi E, Adebiyi M. *et al.* Development of bioinformatics infrastructure for genomics research. *Glob Heart* 2017;**12**:91–8. https://doi.org/10.1016/j.gheart.2017.01.005
35. Mulder N, Schwartz R, Brazas MD. *et al.* The development and application of bioinformatics core competencies to improve bioinformatics training and education. *PLoS Comput Biol* 2018;**14**. https://doi.org/10.1371/journal.pcbi.1005772

36. Aron S, Gurwitz K, Panji S. *et al.* H3abionet: developing sustainable bioinformatics capacity in africa. *EMBnet J* 2017;**23**:886. https://doi.org/10.14806/ej.23.0.886
37. Kibet CK, Entfellner JBD, Jjingo D. *et al.* Designing and delivering bioinformatics project-based learning in East Africa. *BMC Bioinformatics* 2024;**25**:150. https://doi.org/10.1186/s12859-024-05680-2
38. Nanjala R, Nyasimi F, Masiga D. *et al.* A mentorship and incubation program using project-based learning to build a professional bioinformatics pipeline in Kenya. *PLoS Comput Biol* 2023;**19**:e1010904. https://doi.org/10.1371/journal.pcbi.1010904
39. Jjingo D, Mboowa G, Sserwadda I. *et al.* Bioinformatics mentorship in a resource limited setting. *Brief Bioinform* 2022;**23**:bbab399. https://doi.org/10.1093/bib/bbab399
40. Shaffer JG, Mather FJ, Wele M. *et al.* Expanding research capacity in sub-Saharan Africa through informatics, bioinformatics, and data science training programs in Mali. *Front Genet* 2019;**10**:331. https://doi.org/10.3389/fgene.2019.00331
41. Adoga MP, Fatumo SA, Agwale SM. H3Africa: a tipping point for a revolution in bioinformatics, genomics and health research in Africa. *Source Code Biol Med* 2014;**9**:10.
42. Mulder N, Abimiku A, Adebamowo SN. *et al.* H3Africa: current perspectives. *Pharmacogenomics Pers Med* 2018;**11**:59–66. https://doi.org/10.2147/PGPM.S141546
43. Tastan Bishop O, Adebiyi EF, Alzohairy AM. *et al.* Bioinformatics education—perspectives and challenges out of Africa. *Brief Bioinform* 2015;**16**:355–64. https://doi.org/10.1093/bib/bbu022
44. Cloete N, Mouton J, Sheppard C. *Doctoral Education in South Africa (p. 296)*. African Minds, 2016.
45. Korhonen V. *Higher Education Graduation in the US*. Statistica, 2024.
46. Sowell R, Bell N, Francis S. *et al.* The role and status of the master's degree in STEM. *Counc Grad Sch* 2010.
47. Kanyerezi S, Sserwadda I, Ssemaganda A. *et al.* HIV-DRIVES: HIV drug resistance identification, variant evaluation, and surveillance pipeline. *Access Microbiol* 2024;**6**:000815.v3.
48. Ssekagiri A, Jjingo D, Lujumba I. *et al.* QuasiFlow: a Nextflow pipeline for analysis of NGS-based HIV-1 drug resistance data. *Bioinformatics Adv* 2022;**2**:vbac089. https://doi.org/10.1093/bioadv/vbac089
49. Ssekagiri A, Jjingo D, Bbosa N. *et al.* HIVseqDB: a portable resource for NGS and sample metadata integration for HIV-1 drug resistance analysis. *Bioinformatics Adv* 2024;**4**:vbae008. https://doi.org/10.1093/bioadv/vbae008
50. Bbosa N, Ssemwanga D, Nsubuga RN. *et al.* Phylogenetic networks and parameters inferred from HIV nucleotide sequences of high-risk and general population groups in Uganda: implications for epidemic control. *Viruses* 2021;**13**:970. https://doi.org/10.3390/v13060970
51. Kayondo HW, Ssekagiri A, Nabakooza G. *et al.* Employing phylogenetic tree shape statistics to resolve the underlying host population structure. *BMC Bioinformatics* 2021;**22**:546. https://doi.org/10.1186/s12859-021-04465-1
52. Mboowa G, Kakooza F, Egesa M. *et al.* The rise of pathogen genomics in Africa. *F1000Research* 2024;**13**:468. https://doi.org/10.12688/f1000research.147114.2
53. Toure M, Shaffer JG, Sanogo D. *et al.* Seasonal malaria chemoprevention therapy in children up to 9 years of age: protocol for a cluster-randomized trial study. *JMIR Res Protoc* 2024;**13**:e51660. https://doi.org/10.2196/51660
54. Adegbite G, Edeki S, Isewon I *et al.* Investigating the epidemiological factors responsible for malaria transmission dynamics. *InIOP Conference Series: Earth and Environmental Science* 2022 (Vol. 993, No. 1, p. 012008). IOP Publishing.
55. Modeling clinical malaria episodes in different ecological settings in Mali, 2018–2022. *IJID Reg* 2024;**10**:24–30. https://doi.org/10.1016/j.ijregi.2023.11.006
56. Abukari Z, Okonu R, Nyarko SB. *et al.* The diversity, multiplicity of infection and population structure of P. falciparum parasites circulating in asymptomatic carriers living in high and low malaria transmission settings of Ghana. *Genes* 2019;**10**:434. https://doi.org/10.3390/genes10060434
57. Oduselu GO, Afolabi R, Ademuwagun I. *et al.* Structure-based pharmacophore modeling, virtual screening, and molecular dynamics simulation studies for identification of plasmodium falciparum 5-aminolevulinate synthase inhibitors. *Front Med* 2023;**9**. https://doi.org/10.3389/fmed.2022.1022429
58. Oduselu GO, Ajani OO, Ajamma YU. *et al.* Homology modelling and molecular docking studies of selected substituted benzo [d] imidazol-1-yl methyl benzimidamide scaffolds on plasmodium falciparum adenylosuccinate lyase receptor. *Bioinformatics Biol Insights* 2019;**13**:1177932219865533. https://doi.org/10.1177/1177932219865533
59. Afolabi R, Chinedu S, Ajamma Y. *et al.* Computational identification of plasmodium falciparum RNA pseudouridylate synthase as a viable drug target, its physicochemical properties, 3D structure prediction and prediction of potential inhibitors. *Infect Genet Evol* 2022;**97**:105194. https://doi.org/10.1016/j.meegid.2021.105194
60. Afolabi R, Chinedu SN, Adebiyi E. Expression and purification of RNA Pseudouridylate synthase putative of malaria parasite (*Plasmodium falciparum*). *Sci Afr* 2024;**24**:e02160. https://doi.org/10.1016/j.sciaf.2024.e02160
61. Elebiju OF, Oduselu GO, Ogunnupebi TA. *et al.* Design of potential inhibitors of Pf5-ALAS in liver stage plasmodium falciparum: a sustainable chemotherapeutic approach to address antimalarial resistance. *IOP Conf Ser Earth Environ Sci* 2024;**1342**:012006. https://doi.org/10.1088/1755-1315/1342/1/012006
62. Chinedu SN, Bella-Omunagbe M, Okafor E. *et al.* Computational studies on 6-pyruvoyl tetrahydropterin synthase (6-PTPS) of Plasmodium falciparum. *Bioinformatics Biol Insights* 2024;**18**: 11779322241230214. https://doi.org/10.1177/11779322241230214
63. Adebayo GP, Oduselu GO, Aderohunmu DV. *et al.* Structure-based design, and development of amidinyl, amidoximyl and hydroxamic acid based organic molecules as novel antimalarial drug candidates. *Arab J Chem* 2024;**17**:105573. https://doi.org/10.1016/j.arabjc.2023.105573
64. Sangare M, Cissé C, Cruz P. *et al.* Pf phosphatidylserine decarboxylase: molecular modeling and Inhibithors. *J Comput Biol* 2023;**12**:60–78p. https://doi.org/10.37591/rrjocb.v12i2.3318
65. Adedeji EO, Ogunlana OO, Fatumo S. *et al.* Anopheles metabolic proteins in malaria transmission, prevention and control: a review. *Parasit Vectors* 2020;**13**:465. https://doi.org/10.1186/s13071-020-04342-5
66. Oyelade J, Isewon I, Aromolaran O. *et al.* Computational identification of metabolic pathways of plasmodium falciparum using the k-shortest path algorithm. *Int J Genomics* 2019;**2019**:1–13. https://doi.org/10.1155/2019/1750291
67. Okafor A, Adam Y, Brors B. *et al.* Transcriptome analysis reveals a novel DNA element that may interact with chromatin-associated proteins in Plasmodium berghei during erythrocytic development. *Scientific Reports*. 2025;**15**:18621. https://doi.org/10.1101/2024.03.10.584310

68. Traore M, Sangare H, Diabate O. *et al.* Causal effect of severe and non-severe malaria on dyslipidemia in African ancestry individuals: a Mendelian randomization study. *Ann Hum Genet* . 2025;**89**:178–87.
69. Oladejo DO, Anzaku DO, Mamudu CO. *et al.* Gene expression levels and inhibitory effect of 7-[(7-methoxy-4,5-dihydro-1H-benzo[g]indazol-3-yl)carbonyl]-2-phenyl-5,6,7,8-tetrahydropyrazolo[1,5-a]pyrido[4,3-d]Pyrimidin-9(1H)-one (MCL) against AP2-I and BDP1 in malaria experimental models. *Sci Afr* 2024;**23**:e02010.
70. Wodelo W, Wampande EM, Andama A. *et al.* Polymorphisms in immune genes and their association with tuberculosis susceptibility: an analysis of the African population. *Appl Clin Genet* 2024;**17**:33–46. https://doi.org/10.2147/TACG.S457395
71. Kayongo A, Ntayi ML, Olweny G. *et al.* Airway microbiome signature accurately discriminates mycobacterium tuberculosis infection status. *iScience* 2024;**27**:110142. https://doi.org/10.1016/j.isci.2024.110142
72. Kayongo A, Bartolomaeus TUP, Birkner T. *et al.* Sputum microbiome and chronic obstructive pulmonary disease in a rural Ugandan cohort of well-controlled HIV infection. *Microbiol Spectr* 2023;**11**:e0213921. https://doi.org/10.1128/spectrum.02139-21
73. Babirye SR, Nsubuga M, Mboowa G. *et al.* Machine learning-based prediction of antibiotic resistance in mycobacterium tuberculosis clinical isolates from Uganda. *BMC Infect Dis* 2024;**24**:1391. https://doi.org/10.1186/s12879-024-10282-7
74. Obura HO, Mlay CD, Moyo PL. *et al.* Molecular Phylogenetics of HIV-1 subtypes in African populations: a case study of sub-Saharan African countries. *BioRxiv*. 2022:2022-05. https://doi.org/10.1101/2022.05.18.492401
75. Ssemwanga D, Bbosa N, Nsubuga RN. *et al.* The molecular epidemiology and transmission dynamics of HIV type 1 in a general population cohort in Uganda. *Viruses* 2020;**12**:1283. https://doi.org/10.3390/v12111283
76. Bbosa N, Ssemwanga D, Ssekagiri A. *et al.* Phylogenetic and demographic characterization of directed HIV-1 transmission using deep sequences from high-risk and general population cohorts/groups in Uganda. *Viruses* 2020;**12**:331. https://doi.org/10.3390/v12030331
77. Bayigga L, Nabatanzi R, Ssekagiri A. *et al.* Diverse vaginal microbiome was associated with pro-inflammatory vaginal milieu among pregnant women in Uganda. *Hum Microbiome J* 2020;**18**:100076. https://doi.org/10.1016/j.humic.2020.100076
78. Nabisubi P, Kanyerezi S, Kebirungi G. *et al.* Beyond the fever: shotgun metagenomic sequencing of stool unveils pathogenic players in HIV-infected children with non-malarial febrile illness. *BMC Infect Dis* 2025;**25**:96. https://doi.org/10.1186/s12879-025-10517-1
79. COVID-19 pandemic: is Africa different? *J Natl Med Assoc* 2021;**113**:324–35. https://doi.org/10.1016/j.jnma.2020.10.001
80. Koné A, Diallo D, Kané F. *et al.* Dynamics of SARS-CoV-2 variants characterized during different COVID-19 waves in Mali. *IJID Reg* 2023;**6**:24–8. https://doi.org/10.1016/j.ijregi.2022.11.009
81. Bbosa N, Ssemwanga D, Namagembe H. *et al.* Rapid replacement of SARS-CoV-2 variants by delta and subsequent arrival of omicron, Uganda, 2021. *Emerg Infect Dis* 2022;**28**:1021–5. https://doi.org/10.3201/eid2805.220121
82. Laverdure S, Kazadi D, Kone K. *et al.* SARS-CoV-2 seroprevalence in vaccine-naïve participants from the Democratic Republic of Congo, Guinea, Liberia, and Mali. *Int J Infect Dis* 2024;**142**:106985. https://doi.org/10.1016/j.ijid.2024.106985
83. Diabate O, Cisse C, Sangare M. *et al.* Identification of promising high-affinity inhibitors of SARS-CoV-2 main protease from African natural products databases by virtual screening. *Res Sq* 2023;rs.3.rs-2673755. https://doi.org/10.21203/rs.3.rs-2673755/v1
84. Galiwango R, Bainomugisha E, Kivunike F. *et al.* Air pollution and mobility patterns in two Ugandan cities during COVID-19 mobility restrictions suggest the validity of air quality data as a measure for human mobility. *Environ Sci Pollut Res* 2023;**30**:34856–71. https://doi.org/10.1007/s11356-022-24605-1
85. Buyego P, Katwesigye E, Kebirungi G. *et al.* Feasibility of virtual reality based training for optimising COVID-19 case handling in Uganda. *BMC Med Educ* 2022;**22**:274. https://doi.org/10.1186/s12909-022-03294-x
86. Salifu SP, Bukari A-RA, Frangoulidis D. *et al.* Current perspectives on the transmission of Q fever: highlighting the need for a systematic molecular approach for a neglected disease in Africa. *Acta Trop* 2019;**193**:99–105. https://doi.org/10.1016/j.actatropica.2019.02.032
87. Luka MM, Kamau E, de Laurent ZR. *et al.* Whole genome sequencing of two human rhinovirus a types (A101 and A15) detected in Kenya, 2016–2018. *Wellcome Open Res* 2021;**6**:178. https://doi.org/10.12688/wellcomeopenres.16911.2
88. Gregorova M, Santopaolo M, Garner LC. *et al.* Early NK-cell and T-cell dysfunction marks progression to severe dengue in patients with obesity and healthy weight. *Nat Commun* 2025;**16**:5569. https://doi.org/10.1038/s41467-025-60941-9
89. Coulibaly TA, Koné A, Sissoko H. *et al.* Quantification de l'ADN du virus de l'hépatite B: Validation d'une approche de qPCR pour détecter les femmes enceintes à haut risque de transmettre le virus au Mali. *Rev Malienne Infect Microbiol* 2024;**19**:50–5.
90. Namuwulya P, Ashraf S, Niebel M. *et al.* Viruses associated with measles-like illnesses in Uganda. *J Infect* 2024;**88**:106148. https://doi.org/10.1016/j.jinf.2024.106148
91. Fatumo S, Machipisa T, Abdoulaye D. *et al.* GWAS identifies genetic clusters of cardiometabolic risk factors in continental Africans. 2023. https://doi.org/10.21203/rs.3.rs-3458637/v1
92. Udosen B, Soremekun O, Kamiza A. *et al.* Meta-analysis and multivariate GWAS analyses in 80,950 individuals of African ancestry identify novel variants associated with blood pressure traits. *Int J Mol Sci* 2023;**24**:2164. https://doi.org/10.3390/ijms24032164
93. Diawara A, Coulibaly DM, Hussain TYA. *et al.* Type 2 diabetes prevalence, awareness, and risk factors in rural Mali: a cross-sectional study. *Sci Rep* 2023;**13**:3718. https://doi.org/10.1038/s41598-023-29743-1
94. Doumbia L, Findley S, Ba HO. *et al.* Formative research to adapt the 'diabetes prevention program—power to prevent' for implementation in Bamako, Mali. *BMC Health Serv Res* 2024;**24**:1–10.
95. Soremekun C, Jjingo D, Kateete D. *et al.* Structural insights into conformational stability of both wild-type and mutant insulin receptor gene. *Next Res* 2024;**1**:100041.
96. Diawara A, Traore M, Diabaté O. *et al.* Genetically proxied therapeutic inhibition of kidney function drug targets and type 2 diabetes in Africans: a Mendelian randomization study. *Sci Prog* 2025;**108**:00368504251338631. https://doi.org/10.1177/00368504251338631
97. Diawara A, Coulibaly DM, Kone D. *et al.* Dyslipidemia in adults with type 2 diabetes in a rural community in Ganadougou, Mali: a cross-sectional study. *J Diabetes Mellitus* 2024;**14**:133–52. https://doi.org/10.4236/jdm.2024.142012
98. Kamiza AB, Touré SM, Zhou F. *et al.* Multi-trait discovery and fine-mapping of lipid loci in 125,000 individuals of African ancestry. *Nat*

Commun 2023;**14**:5403. https://doi.org/10.1038/s41467-023-41271-0

99. Mayanja R, Kintu C, Diabate O. *et al.* Molecular dynamic simulation reveals structure differences in APOL1 variants and implication in pathogenesis of chronic kidney disease. *Genes* 2022;**13**:1460. https://doi.org/10.3390/genes13081460
100. Ademuwagun IA, Oduselu GO, Rotimi SO. *et al.* Pharmacophore-aided virtual screening and molecular dynamics simulation identifies TrkB agonists for treatment of CDKL5-deficiency disorders. *Bioinformatics Biol Insights* 2023;**17**:11779322231158254. https://doi.org/10.1177/11779322231158254
101. Nsubuga M, Mutegeki H, Jjingo D. *et al.* The Ugandan sickle pan-African research consortium registry: design, development, and lessons. *BMC Med Inform Decis Mak* 2024;**24**:212.
102. McCormack VA, Schüz J. Africa's growing cancer burden: environmental and occupational contributions. *Cancer Epidemiol* 2012;**36**:1–7. https://doi.org/10.1016/j.canep.2011.09.005
103. Hamdi Y, Abdeljaoued-Tej I, Zatchi AA. *et al.* Cancer in Africa: the untold story. *Front Oncol* 2021;**11**:650117. https://doi.org/10.3389/fonc.2021.650117
104. Ngwa W, Addai BW, Adewole I. *et al.* Cancer in sub-Saharan Africa: a lancet oncology commission. *Lancet Oncol* 2022;**23**:e251–312. https://doi.org/10.1016/S1470-2045(21)00720-8
105. Duduyemi BM, Agyemang DL, Adankwah E. *et al.* Knowledge, perception and screening of local dye workers regarding urinary bladder cancer in Ghana. *Afr J Urol* 2020;**26**:59. https://doi.org/10.1186/s12301-020-00074-1
106. Oppong R, Yeboah D, Owusu-Ansah M. *et al.* Radiation therapy for breast cancer in Africa. *Adv Radiat Oncol* 2024;**9**:101488. https://doi.org/10.1016/j.adro.2024.101488
107. Appiah EM, Yakubu B, Salifu SP. Comprehensive microbial network analysis of gastric microbiome reveal key species affecting gastric carcinogenesis. *Microbe* 2023;**1**:100009. https://doi.org/10.1016/j.microb.2023.100009
108. Lubwama M, Kateete DP, Katende G. *et al.* CTX-M, TEM, and SHV genes in Escherichia coli, Klebsiella pneumoniae, and Enterobacter spp isolated from hematologic cancer patients with bacteremia in Uganda. *Infect Drug Resist* 2024;**17**:641–53. https://doi.org/10.2147/IDR.S442646
109. Doughan A, Salifu SP. Genes associated with diagnosis and prognosis of Burkitt lymphoma. *IET Syst Biol* 2022;**16**:220–9. https://doi.org/10.1049/syb2.12054
110. Chen Y, Fachko D, Ivanov NS. *et al.* B cell receptor-responsive miR-141 and viral miR-BART9 promote Epstein—Barr virus reactivation through FOXO3 inhibition. *bioRxiv*. 2019:837294. https://doi.org/10.1101/837294
111. Salifu SP, Doughan A. New clues to prognostic biomarkers of four hematological malignancies. *J Cancer* 2022;**13**:2490–503. https://doi.org/10.7150/jca.69274
112. Nabweyambo S, Sande OJ, McGovern N. *et al.* Circulating levels of angiogenic factors and their association with preeclampsia among pregnant women at Mulago National Referral Hospital in Uganda. *PloS One* 2021;**16**:e0251227. https://doi.org/10.1371/journal.pone.0251227
113. Sory PI, Sidi T, Kodio B. *et al.* Fréquence hospitalière de l'ostéoporose chez la femme ménopausée dans le service de rhumatologie au CHU point G. à Bamako. *Rhumatol Afr Francoph* 2024;**7**:11–5.
114. Ademuwagun IA, Rotimi SO, Syrbe S. *et al.* Voltage gated sodium channel genes in epilepsy: mutations, functional studies, and treatment dimensions. *Front Neurol* 2021;**12**.
115. Linda L, Mutema AB, Babirye SR. *et al.* Multi-ancestry analysis of plasma protein levels influencing and responding to major depression liability. 2025. https://doi.org/10.21203/rs.3.rs-5828682/v1
116. Mensah I, Amoako-Yirenkyi P, Frempong NK. *et al.* Wavelets based feature extraction with PCA for predicting autism In neonates using Navïe Bayes classifier. 2021. https://doi.org/10.21203/rs.3.rs-1048775/v1
117. Soremekun O, Musanabaganwa C, Uwineza A. *et al.* A Mendelian randomization study of genetic liability to post-traumatic stress disorder and risk of ischemic stroke. *Transl Psychiatry* 2023;**13**:1–6.
118. Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet Lond Engl* 2022;**399**:629–55.
119. Oduselu GO, Aderohunmu DV, Ajani OO. *et al.* Synthesis, in silico and in vitro antimicrobial efficacy of substituted arylidene-based quinazolin-4(3H)-one motifs. *Front Chem* 2023;**11**. https://doi.org/10.3389/fchem.2023.1264824
120. Carolus H, Van Dyck K, Van Dijck P. Candida albicans and staphylococcus species: a threatening twosome. *Front Microbiol* 2019;**10**:2162. https://doi.org/10.3389/fmicb.2019.02162
121. Kalambry AC, Potindji TMF, Guindo I. *et al.* Resistance phenotypes and molecular characteristics of *Staphylococcus aureus* associated with pleuritis in patients at "Hôpital du Mali" teaching hospital. *Research Square*. 2024:rs-3. https://doi.org/10.21203/rs.3.rs-3579825/v1
122. Workneh M, Hamill MM, Kakooza F. *et al.* Antimicrobial resistance of Neisseria Gonorrhoeae in a newly implemented surveillance program in Uganda: surveillance report. *JMIR Public Health Surveill* 2020;**6**:e17009. https://doi.org/10.2196/17009
123. Kakooza F, Musinguzi P, Workneh M. *et al.* Implementation of a standardised and quality-assured enhanced gonococcal antimicrobial surveillance programme in accordance with WHO protocols in Kampala, Uganda. *Sex Transm Infect* 2021;**97**:312–6. https://doi.org/10.1136/sextrans-2020-054581
124. Melendez J, Emmanuel M, Aketoko AO. *et al.* Characterization of pharyngeal gonorrhea in Ugandan men with urethral discharge syndrome. *Sex Transm Infect* 2021;**97**:A98–9.
125. Trick AY, Melendez JH, Chen FE. *et al.* A portable magnetofluidic platform for detecting sexually transmitted infections and antimicrobial susceptibility. *Sci Transl Med* 2021;**13**:eabf6356. https://doi.org/10.1126/scitranslmed.abf6356
126. Mayito J, Tumwine C, Galiwango R. *et al.* Combating antimicrobial resistance through a data-driven approach to optimize antibiotic use and improve patient outcomes: protocol for a mixed methods study. *JMIR Res Protoc* 2024;**13**:e58116.
127. Ogunnupebi TA, Oduselu GO, Elebiju OF. *et al.* In silico studies of benzimidazole derivatives as sustainable inhibitors against methicillin-resistant Staphylococcus aureus. *IOP Conf Ser Earth Environ Sci* 2024;**1342**:012005. https://doi.org/10.1088/1755-1315/1342/1/012005
128. Elebiju OF, Oduselu GO, Ogunnupebi TA. *et al.* In silico design of potential small-molecule antibiotic adjuvants against salmonella typhimurium ortho acetyl sulphydrylase synthase to address antimicrobial resistance. *Pharmaceuticals* 2024;**17**:543. https://doi.org/10.3390/ph17050543
129. Aderohunmu DV, Ajani OO, Oduselu GO. *et al.* Microwave-assisted synthesis of coumarin-based 5,6-dihydro pyrimidin-2(1H)-one derivatives. *J Phys Conf Ser* 2019;**1299**:012119. https://doi.org/10.1088/1742-6596/1299/1/012119

130. Oyelade J, Isewon I, Uwoghiren E. *et al.* In Silico knockout screening of plasmodium falciparum reactions and prediction of novel essential reactions by analysing the metabolic network. *Biomed Res Int* 2018;**2018**:1–11. https://doi.org/10.1155/2018/8985718
131. Aromolaran O, Beder T, Adedeji E. *et al.* Predicting host dependency factors of pathogens in *Drosophila melanogaster* using machine learning. *Comput Struct Biotechnol J* 2021;**19**:4581–92. https://doi.org/10.1016/j.csbj.2021.08.010
132. Aromolaran O, Beder T, Oswald M. *et al.* Essential gene prediction in *Drosophila melanogaster* using machine learning approaches based on sequence and functional features. *Comput Struct Biotechnol J* 2020;**18**:612–21. https://doi.org/10.1016/j.csbj.2020.02.022
133. Beder T, Aromolaran O, Dönitz J. *et al.* Identifying essential genes across eukaryotes by machine learning. *NAR Genomics Bioinforma* 2021;**3**:lqab110. https://doi.org/10.1093/nargab/lqab110
134. Kapalaga G, Kivunike FN, Kerfua S. *et al.* Enhancing random forest predictive performance for foot and mouth disease outbreaks in Uganda: a calibrated uncertainty prediction approach for varying distributions. *Front Artif Intell* 2024;**7**:1455331. https://doi.org/10.3389/frai.2024.1455331
135. Aromolaran O, Aromolaran D, Isewon I. *et al.* Machine learning approach to gene essentiality prediction: a review. *Brief Bioinform* 2021;**22**:bbab128. https://doi.org/10.1093/bib/bbab128
136. Adegbite G, Edeki S, Isewon I. *et al.* Mathematical modeling of malaria transmission dynamics in humans with mobility and control states. *Infect Dis Model* 2023;**8**:1015–31. https://doi.org/10.1016/j.idm.2023.08.005
137. Oyelade J I Isewon, O Oladipupo. *et al.* Data clustering: algorithms and its applications. in *2019 19th International Conference on Computational Science and Its Applications (ICCSA)* 71–81 (2019). https://doi.org/10.1109/ICCSA.2019.000-1.
138. Nanjala R, Mbiyavanga M, Hashim S. *et al.* Assessing HLA imputation accuracy in a west African population. *PloS One* 2023;**18**:e0291437. https://doi.org/10.1371/journal.pone.0291437
139. Soyemi J, Isewon I, Oyelade J. *et al.* Inter-species/host-parasite protein interaction predictions reviewed. *Curr Bioinformatics* 2018;**13**:396–406.
140. Kariuki EG, Kibet C, Paredes JC. *et al.* Metatranscriptomic analysis of the gut microbiome of black soldier fly larvae reared on lignocellulose-rich fiber diets unveils key lignocellulolytic enzymes. *Front Microbiol* 2023;**14**:1120224. https://doi.org/10.3389/fmicb.2023.1120224
141. Duduyemi BM, Andoh D, Adankwah E. *et al.* The use of special stains in the detection of vascular invasion in cases of colon cancer in resource-limited settings in Africa. *Ann Trop Pathol* 2020;**11**:21. https://doi.org/10.4103/atp.atp_1_20
142. Borkor RN, Svärd M, Amoako-Yirenkyi P. A stable scheme of the curvilinear shallow water equations with no-penetration and far-field boundary conditions. *Comput Fluids* 2024;**269**:106136. https://doi.org/10.1016/j.compfluid.2023.106136
143. Schwartz R, Brooksbank C, Gaeta B. *et al.* Guidelines for developing and updating short courses and course programs using the ISCB competency framewor. 2021. https://doi.org/10.5281/zenodo.5418103
144. Fatumo S, Ebenezer TGE, Ekenna C. *et al.* The Nigerian bioinformatics and genomics network (NBGN): a collaborative platform to advance bioinformatics and genomics in Nigeria. *Glob Health Epidemiol Genomics* 2020;**5**:e3. https://doi.org/10.1017/gheg.2020.3
145. Akurugu WA, Doughan A, Adamu Bukari AR. *et al.* Highlights of the 3rd ISCB Africa student council symposium 2019 in Ghana. *F1000Res* 2020;**9**. https://doi.org/10.12688/f1000research.24101.1
146. Fadlelmola FM. *et al.* H3ABioNet genomic medicine and microbiome data portals hackathon proceedings. *Database J Biol Databases Curation* 2021;**2021**:baab016.
147. Giovanni MY. African centers of excellence in bioinformatics and data intensive science: building capacity for enhancing data intensive infectious diseases research in Africa. *Journal of infectious diseases & microbiology.* 2023;**1**:006. https://doi.org/10.37191/Mapsci-JIDM-1(2)-006
148. Hurt DE, Whalen C, Wele M. *et al.* African centers of excellence in bioinformatics: an evidence-based approach to biomedical research collaboration in Africa. *Am J Trop Med Hyg* 2019;**101**:224–4.
149. Please H, Narang K, Bolton W. *et al.* Virtual reality technology for surgical learning: qualitative outcomes of the first virtual reality training course for emergency and essential surgery delivered by a UK–Uganda partnership. *BMJ Open Qual* 2024;**13**:e002477. https://doi.org/10.1136/bmjoq-2023-002477
150. McMullen E, Kamurari S, Price R. *et al.* Emerging technology solutions to support national emergency workforce capacity-building initiatives: lessons from Ugandan policy and practice. *Emerg Med J* 2025;**42**:62–9.
151. Shaw J, Ali J, Atuire CA. *et al.* Research ethics and artificial intelligence for global health: perspectives from the global forum on bioethics in research. *BMC Med Ethics* 2024;**25**:46. https://doi.org/10.1186/s12910-024-01044-w
152. Marvin G, Tamale M, Kanagwa B. *et al.* Philosophical review of artificial intelligence for society 5.0. in *Proceedings of International Conference on Paradigms of Communication, Computing and Data Analytics* (eds Yadav, A., Nanda, S. J. & Lim, M.-H.) 1–15 (Springer Nature, Singapore, 2023). https://doi.org/10.1007/978-981-99-4626-6_1.
153. Organization, W. H. *Ethics and Governance of Artificial Intelligence for Health: Large Multi-Modal Models. WHO Guidance*. World Health Organization, 2024.
154. Ferretti A, Adjei KK, Ali J. *et al.* Digital tools for youth health promotion: principles, policies and practices in sub-Saharan Africa. *Health Promot Int* 2024;**39**:daae030.
155. Ghazal H, Adam Y, Idrissi Azami A. *et al.* Plant genomics in Africa: present and prospects. *Plant J* 2021;**107**:21–36. https://doi.org/10.1111/tpj.15272
156. Doughan A, Adingo W, Salifu SP. RNA-seq research landscape in Africa: systematic review reveals disparities and opportunities. *Eur J Med Res* 2023;**28**:244. https://doi.org/10.1186/s40001-023-01206-3
157. Kwizera R, Mande E, Omali D. *et al.* Translational research in Uganda: linking basic science to bedside medicine in a resource limited setting. *J Transl Med* 2021;**19**:76. https://doi.org/10.1186/s12967-021-02747-z
158. NIH RePORTER. https://reporter.nih.gov/.
159. Onywera H, Ondoa P, Nfii F. *et al.* Boosting pathogen genomics and bioinformatics workforce in Africa. *Lancet Infect Dis* 2024;**24**:e106–12. https://doi.org/10.1016/S1473-3099(23)00394-8
160. Casad BJ, Franks JE, Garasky CE. *et al.* Gender inequality in academia: problems and solutions for women faculty in STEM. *J Neurosci Res* 2021;**99**:13–23. https://doi.org/10.1002/jnr.24631

161. Chavatzia T. *Cracking the Code: Girls' and women's Education in Science, Technology, Engineering and Mathematics (STEM)*. Paris Fr. U. N. Educ. Sci. Cult. Organ, 2017.
162. Founou LL, Yamba K, Kouamou V. *et al.* African women in science and development, bridging the gender gap. *World Dev Perspect* 2023;**31**:100528. https://doi.org/10.1016/j.wdp.2023.100528
163. Galiwango R, Whalen CJ, Kebirungi G *et al.* Ten simple rules for building and maintaining sustainable high-performance computing infrastructure for research in resource-limited settings. *PLOS Computational Biology* 2025;**21**:e1013481. https://doi.org/10.1371/journal.pcbi.1013481

Briefings in Bioinformatics, 2026, **27**, bbag046
https://doi.org/10.1093/bib/bbag046
Review