Career feature

https://doi.org/10.1038/s41587-023-01933-2

Bridging the gap in African biodiversity genomics and bioinformatics

The Open Institute of the African BioGenome Project empowers African scientists and institutions with the skill sets, capacity and infrastructure to advance scientific knowledge and innovation and drive economic growth.

Africa, a continent of 1.3 billion people, had 326 researchers per million people in 2018[1,2], whereas the global average for the number of researchers per million people is 1,368[1,2]. Nevertheless, a strong research community is a requirement to advance scientific knowledge and innovation and drive economic growth[3,4]. This low number of researchers extends to scientific research across Africa and applies to genomic projects such as the African BioGenome Project (AfricaBP) (https://africanbiogenome.org/)[5].

The AfricaBP plans to sequence 100,000 endemic African species in 10 years[5] with an estimated 203,000 gigabases of DNA sequence. AfricaBP aims to generate and analyze these genomes on the ground in Africa. However, in order for the project to achieve these goals, there is a need to empower African scientists and institutions to obtain the required skill sets, capacity and infrastructure to generate, analyze and utilize these sequenced genomes in-country.

The Open Institute is the genomics and bioinformatics knowledge exchange program for the AfricaBP (Figs. 1 and 2). It consists of 10 participating institutions, including the University of South Africa in South Africa and the National Institute of Agricultural Research (INRA) in Morocco. It aims to develop biodiversity genomics and bioinformatics curricula targeted at African scientists; promote and develop genomics and bioinformatics tools that will address critical needs relevant to the African terrain, such as limited internet access; and advance grassroot knowledge exchange through outreach and public engagement, such as quarterly training and workshops.

The AfricaBP Open Institute is designed to close infrastructural gaps that exist in the biodiversity genomics space and build a critical mass of researchers across Africa. For instance, the sequencing of the genome of hyacinth bean (*Lablab purpureus*), the first chromosome-scale plant genome assembly locally sequenced in Africa, benefited from in-depth bioinformatics training of four African scientists in African yam bean (*Sphenostylis stenocarpa*) over a period of eight months[6]. Using this model, assuming other factors such as funding remain constant, that all 100,000 African endemic species are sequenced and analyzed through the AfricaBP Open Institute over the next 10 years and that two genomes are analyzed per four African scientists, we estimate that the sequencing project could train 200,000 African scientists in genomics and bioinformatics.

To deliver these goals, the AfricaBP Open Institute has established openly accessible workshops. The AfricaBP Open Institute workshop on endemic African species was held in May 2022 with nearly 300 registered attendees from across 20 African institutions and 29 countries. This workshop showcased the sequencing and assembly of two African endemic species in partnership with Inqaba Biotechnical Industries and the Vertebrate Genome Project (VGP). The AfricaBP Open Institute workshop on biodiversity genomic technologies and infrastructures was held in September 2022 with more than 400 registered participants from across 28 African countries. This showcased the cutting-edge technologies shaping the biodiversity genomics field, including current understanding on global genomic databases, tools and resources.

Here, we discuss these workshops and demonstrate how the AfricaBP Open Institute efforts could be further developed through a distributed hub-and-spoke model (Figs. 1 and 2).

Five priorities

The Open Institute of the AfricaBP aims to lower some of the barriers that prevent the advancement of biodiversity genomics and bioinformatics knowledge exchange across Africa. It has five critical priority areas, which are discussed below.

Curriculum development. One barrier in the development of training materials in Africa is training that is given in languages that some participants may not fully understand[7,8]. For example, genomic and bioinformatics courses are held in English for participants in Egypt, where most people speak Arabic[9,10]. The AfricaBP Open Institute aims to ensure that biodiversity genomic data and resources follow the FAIR principles (Findable, Accessible, Interoperable and Reusable). However, with the current systems of training, genomics data cannot adhere to the FAIR principles if the data are only findable in English or French. For example, an Egyptian scientist who understands Arabic better than English will not be able to maximize uptake and assimilation of training materials if such materials are only available in English or French[11].

Learning in local languages, such as Swahili, would glue down knowledge learned in the participant's native environment, enable a deeper understanding of concepts and help link biodiversity genomic practices to local knowledge[12–15]. For instance, in a focused group discussion among 115 participants from Yoruba populations in West Africa, it was evident that words that describe the heritability of characters, traits and diseases, such as horses inheriting the ability to race (*ere sisa la fi bi eshin*) and the heritability of sickle cell disease (*arunmolegun*), exist in the Yoruba language; this can be used to improve the understanding of prior informed consent in genomics research[16].

To address this challenge and advance the FAIR principles, the AfricaBP Open Institute will promote the incorporation of local languages in prior informed consents and the use of machine learning and natural language

processing tools for translating training materials into selected widely spoken African languages, such as Kikongo, Swahili, isiZulu, Wolof, Arabic and Amharic. This will be done by partnering with organizations such as Masakhane, an African organization whose mission is to strengthen natural language processing research in African languages[17].

Another barrier to curriculum development in Africa is the use of data from non-African species as training datasets. Participants will benefit from training exercises that are carried out with species endemic to the country where the training is taking place. For instance, in West Africa, sequenced genomes of African locust bean (*Parkia biglobosa*) and bush mango (*Cordyla pinnata*) could be used for training students in structural genomics to understand genome organization and easily link gene content, adaptability and socio-economic relevance. Second, the classical organisms used for bioinformatics trainings in Africa, for example *Escherichia coli* or *Homo sapiens*[18,19], do not provide a broad repertoire of the challenges—such as difficulty in sequencing or high genetic diversity—that exist in working with non-model, endemic species. For instance, non-model vertebrates such as African lions and leopards will require unconventional sample materials and sequencing approaches[20–25].

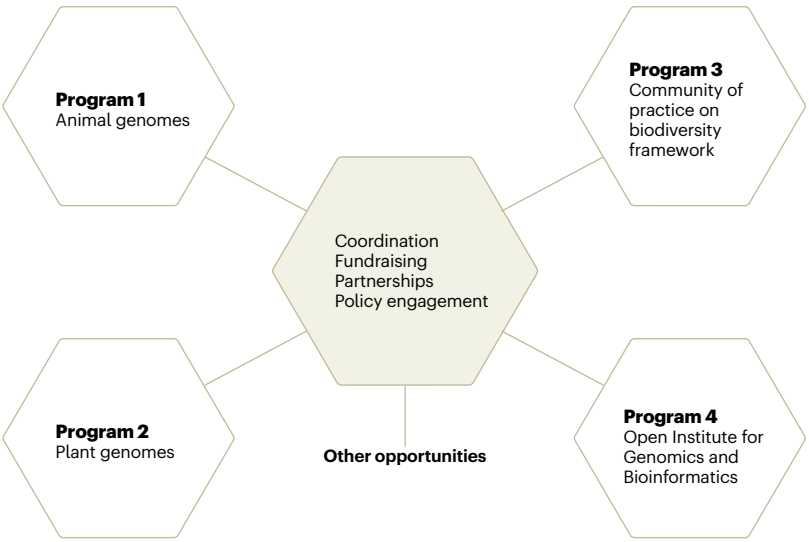

Fig. 1 | The African BioGenome Project will leverage four programs to achieve its goals. Program 1, also called the Nelson Mandela Genomes Initiative for Conservation of Nature, focuses on animal species. Program 2 focuses on plant species. Program 3 focuses on issues around Access and Benefit Sharing of the Nagoya Protocol and the Post-2020 Global Biodiversity Framework. Program 4 is the AfricaBP Open Institute for Genomics and Bioinformatics.

Further develop technology and infrastructure. To enable knowledge exchange between the human genetics and biodiversity genomics communities across Africa, the AfricaBP Open Institute will work with communities such as the bioinformatics network for H3Africa (H3ABioNet) and other African human health and agricultural consortia to advance conversations toward a centralized bioinformatics infrastructure in Africa (Fig. 2). This could be similar to the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) for the European continent. Such infrastructure will provide expertise, services and coordination for genomics and bioinformatics in data generation, storage, sharing and access to genomic data across Africa. Such infrastructure will also increase collaborations between, and maximize resources for, the African human genetics community and the African biodiversity genomics community (Fig. 2). For example, a research group within the African human genetics community that develops a bioinformatics tool could easily share this tool with another research group in the African biodiversity genomics community. The genomics arm of the AfricaBP Open Institute will form specialized African Centres of Excellence for Biodiversity Genomics, employing a public–private partnership with a distributed hub-and-spoke model (Fig. 2). This could be similar to the African Centres of Excellence in infectious diseases in Mali, Nigeria and Uganda[26,27].

Similarly, one of the challenges experienced so far in the AfricaBP pilot project is the logistics of transporting samples from field locations to sequencing centers. The AfricaBP Open Institute aims to develop technology and tools to coordinate the logistics of sample collection and transport from sampling locations to sequencing centers using modular, lightweight supply chain technology.

Promote grassroots knowledge exchange and equitable partnerships. A major goal of the AfricaBP Open Institute is to build grassroots capacity in regard to digital sequence information (DSI), Nagoya Protocol procedures and the post-2020 global biodiversity framework[5]. Bioinformatics and genomics capacity and infrastructure are not equally distributed across Africa, which could result in differences between countries in their ability to exploit the resulting sequencing data and resources. Under-resourced groups across Africa are unlikely to have all of these genomic analytical skills. Hence, they will require support to ensure they are able to gain from benefits derived[5]. The AfricaBP Open Institute aims to support African research and academic institutions with fewer resources by facilitating data analysis and building open tools and software to support their research needs. For instance, the AfricaBP has already engaged with communities such as Galaxy Africa to collaborate on ensuring these resources are available to the AfricaBP community.

The AfricaBP Open Institute also aims to support the formation of equitable partnerships through co-creation of project proposals, joint mobilization of resources between African scientists and their partners, delivery of short- and longer-term courses, workshops and wet-lab trainings, and online and in-person residential capacity-building programs to facilitate knowledge exchange.

Maximize data ownership and sovereignty. Three entities across Africa will particularly benefit from AfricaBP: Africa's early-career researchers and established scientists, agricultural and biodiversity conservation centers, and policy makers such as those designated as Africa's National Focal Points for the Global Environment Facility and Access and Benefit Sharing of the Nagoya Protocol within the structures of the Convention on Biological Diversity (CBD).

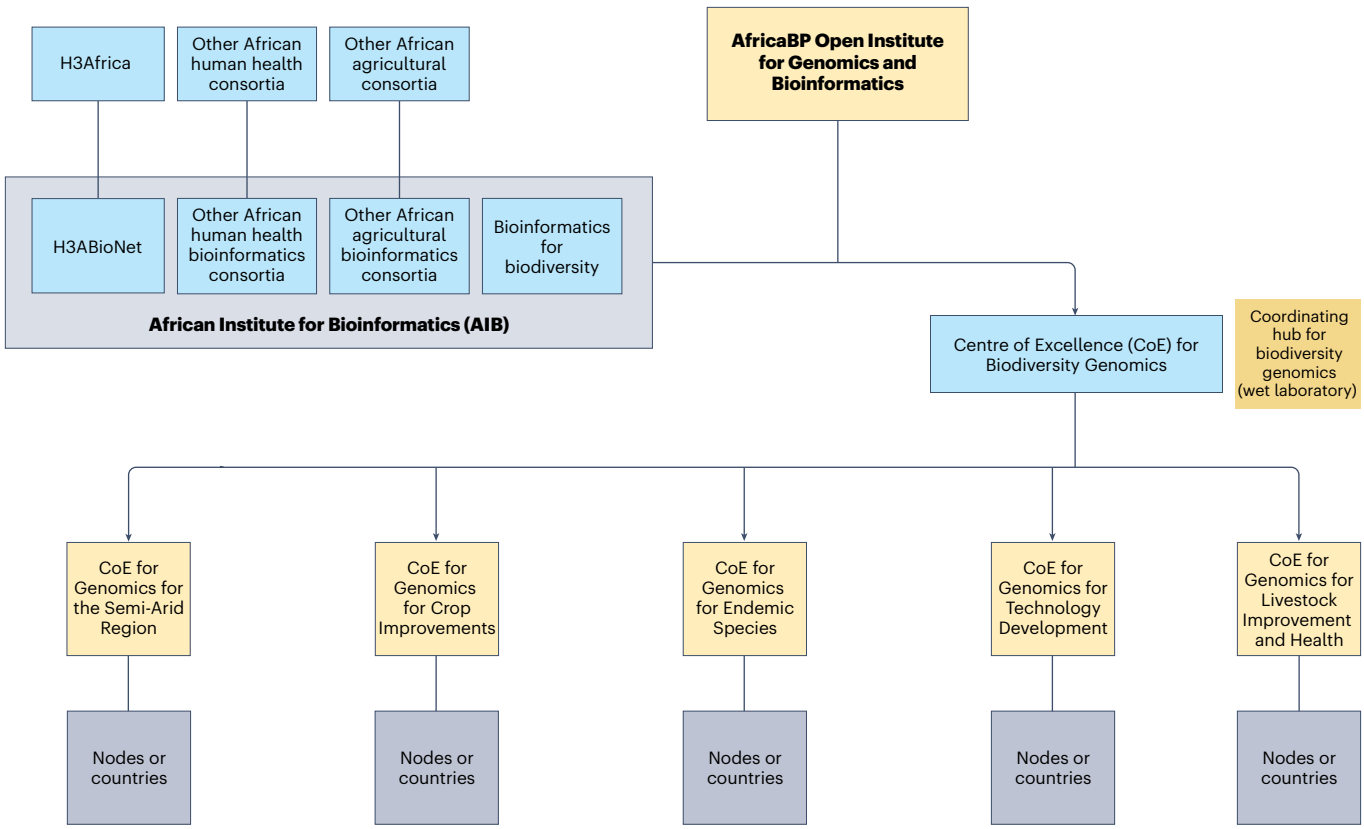

Fig. 2 | The AfricaBP Open Institute will operate a distributed model. The bioinformatics arm of the AfricaBP Open Institute (Bioinformatics for Biodiversity) will link up with other bioinformatics consortia, such as the H3ABioNet and the Bioinformatics Community of Practice, across Africa to set up the African Institute for Bioinformatics (AIB). Bottom boxes depict regional hubs for specific areas of expertise for the Center of Excellence (CoE) linked to participating African countries or nodes. H3ABioNet is the bioinformatics network for H3Africa. Other African human health and agricultural bioinformatics consortia are independent communities and projects that are not part of the AfricaBP Open Institute.

In 2022 the CBD, the United Nations arm responsible for issues concerning biodiversity, adopted the Post-2020 Global Biodiversity Framework. This is a set of principles aimed to safeguard at least 90% of global biodiversity by 2030. At the moment, most African countries and entities described above are limited by instruments and by their capacities to maximize the benefits that the Post-2020 Global Biodiversity Framework presents, as well as by concerns around data ownership and governance.

The AfricaBP Open Institute presents the opportunity for African countries to build the capacities required to benefit from the Post-2020 Global Biodiversity Framework. For instance, an AfricaBP Open Institute workshop is planned for late 2023 to create awareness on issues around DSI, the Post-2020 Biodiversity Framework and the Nagoya Protocol. Subsequent workshops will include training to equip Africa's policy makers, such as the National Focal Points of the CBD, to be able to take advantage of this.

Encourage scientific entrepreneurship and industry. In 2019 the African Continental Free Trade Area (AfCFTA) agreement, the largest single market in the world by number of participating countries, came into force to connect 1.3 billion people across 55 countries with a combined gross domestic product of US$3.4 trillion[28]. Among AfCFTA's many benefits, it will enable and increase intra-Africa trade and the consumption of African bioeconomy products such as agricultural products and services[29,30].

The AfricaBP Open Institute will promote scientific entrepreneurship to support Africa's bioeconomy. For instance, African countries experience higher operating costs for science laboratories compared to others because most consumables, such as reagents and equipment, are imported to Africa from Europe, North America or Asia. The added shipping costs plus the levies and duties imposed by the importing countries drive the costs up to make African laboratories non-competitive with research laboratories in developed or industrialized countries[31,32]. The AfricaBP Open Institute will engage in conversations with African policy makers and governments to provide exemptions from duties on certified scientific consumables and equipment and will support scientists and entrepreneurs in leveraging the AfCFTA to lower laboratory operational costs, for example by setting up franchises for local manufacturing of reagents and equipment, laboratory sharing and hackathons[33,34].

Similarly, insufficient capacity to translate genomic research into commercial products is a major challenge for Africa's research ecosystem. For example, the diversity of some South African endophytes, bacterial or fungal species that live within a plant and have an endosymbiotic relationship, has been established.

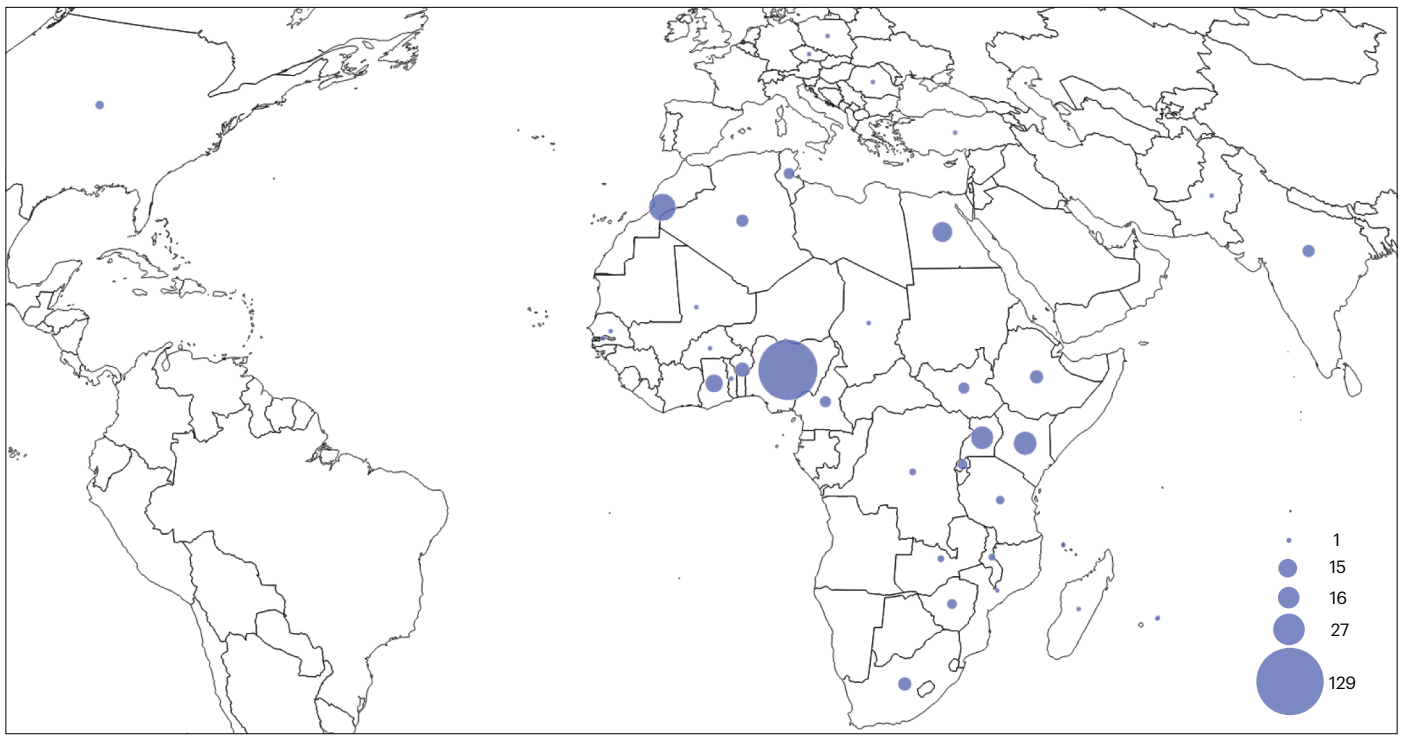

Fig. 3 | Participant representation for the first AfricaBP Open Institute workshop on endemic African species. All geographic regions of Africa were represented, as well as non-African countries. The sizes of the circles indicate the number of participants from each country.

However, the biotechnology potential of these species has not been fully explored[35]. The AfricaBP Open Institute will promote the use and translation of findings from basic genomic sequencing efforts into societal benefits. For example, an AfricaBP scientist who identifies a bioactive metabolite from the yew tree (*Taxus brevifolia*) in South Africa could partner with an AfricaBP industry partner to commercialize it in compliance with the South African laws on Access and Benefit Sharing[36].

Moving forward with purpose

The AfricaBP Open Institute kickstarted the process of implementing its priorities through workshops. In May 2022 it organized the First AfricaBP Open Institute workshop on endemic African species, and in September 2022 it organized the AfricaBP Open Institute workshop on biodiversity genomic technologies and infrastructure.

The First AfricaBP Open Institute workshop on endemic African species recorded a total of 306 applications, 292 of them from 29 African countries, including Nigeria, Morocco and Uganda (Fig. 3). The applicants were affiliated with 193 African organizations and had various educational backgrounds, ranging from graduate students to full professors. The two-day workshop involved theoretical presentations and demonstrations of the journey undertaken during the sequencing of the first two AfricaBP genomes (the speckled mousebird, *Colius striatus*, and beaked blind snake, *Rhinotyphlops lalandei*) in collaboration with Inqaba Biotechnical Industries and the VGP, and included sample collections and processing, sample permits acquisition, ethical considerations, library preparations and sequencing, quality control, assembly and assembly reproduction.

In addition, apart from theoretical presentations, the workshop also undertook a four-week remote residential practical exercise focusing on genome assembly reproduction (via the Slack instant messaging platform) to train selected African scientists in genome assembly using the Galaxy Europe instance (https://assembly.usegalaxy.eu) in collaboration with the VGP (Table 1). Step-by-step practical demonstrations and hands-on tutorials were provided on the VGP pipeline to assemble sample data from the yeast (*Saccharomyces cerevisiae* S288C) genome. Afterwards, the genome assembly of *C. striatus* was reproduced using the VGP pipeline[37].

The aim of the four-week residential post-workshop exercise was to reproduce the assembly of the *C. striatus* genome, ensuring that the results are comparable to those generated by the VGP team (Table 1). The training began with the acquisition of the most recent VGP workflow from the official GitHub repository (https://tinyurl.com/vgpassemblyv2). The workflow files were imported to the Galaxy Europe platform as described on the official VGP GitHub repository. The VGP pipeline consists of five main workflows: Meryldb creation, Hifiasm-HiC-assembly, Purged-assembly, Bionano scaffolding and Hi-C scaffolding[37]. The Purged-assembly and Bionano scaffolding workflows are optional, with the former being used to purge duplications and being necessary only when duplications are identified by quality control after the assembly workflow[37].

The AfricaBP Open Institute workshop on biodiversity genomic technologies and infrastructure was attended by over 400 individuals from 28 African countries. More than 70% of those registered were early-career researchers with ongoing or upcoming projects in diverse areas of genomics (such as plant, vertebrate and microbial genomics). Fifty-six percent of the applicants were very

Table 1 | Genome statistics for the assembly of the *Colius striatus* produced by the VGP assembly team versus the assembly reproduced by AfricaBP Open Institute selected attendees

Evaluating metrics	Scaffolds[a]	Assembly category	VGP assembly	AfricaBP assembly
Total length (bp)	Post-scaffolding	Scaffold assembly	1,159,734,284	1,159,722,284
Number of contigs/scaffolds			227	219
BUSCO score[b]			C:3009[S:2987,D:22], F:42,M:303,n:3354	C:3008[S:2987,D:21],F:42,M:304,:n:3354
Scaffold N50 (bp)			58,796,297	58,796,297

[a]Only post-scaffolding is reported. [b]Complete (C) [single-copy (S), duplicated (D)], fragmented (F), missing (M).

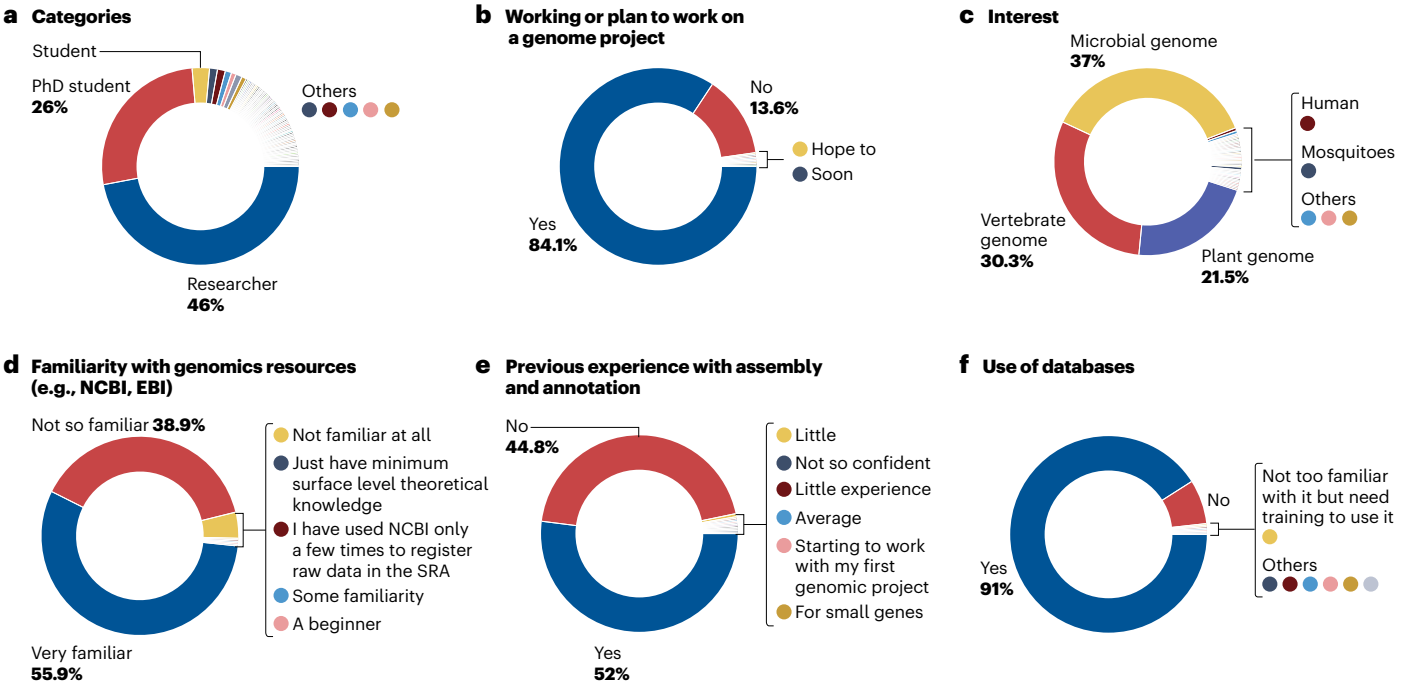

Fig. 4 | Response of pre-workshop survey of the AfricaBP workshop on genomic technologies and infrastructures to identify skill needs of African researchers and genomic gaps in Africa. Participants' responses with respect to familiarity with genomic resources/databases and past experience in genome assembly and annotation. Number of respondents, 404.

familiar with genomics resources hosted at the National Center for Biotechnology Information (NCBI) and European Bioinformatics Institute (EMBL-EBI). About 45% had limited experiences with genome assembly and annotation of genomes, while 91% had prior experience with global genomic databases (Fig. 4).

Feedback from participants showed that at least 95% described the workshop as interesting and felt it presented the right amount of information (Fig. 5). Participants also found that the workshop provided sufficient time for discussion, met their expectations and was of the right length of time. Importantly, 98% of participants expressed interest in attending future workshops of similar nature that AfricaBP will organize (Fig. 5).

Conclusions and next steps

The organization of the AfricaBP Open Institute workshop on endemic African species and biodiversity genomics technologies and infrastructure presents clear indicators for training African scientists in genomic procedures, technologies and infrastructures as well as in the ethical, legal and social issues that accompany genomics practices. This is reflected in the increasing number of participants (Figs. 3, 4 and 5) in the workshops, the involvement of African scientists in diverse genomic projects, and participants' satisfaction with material and content delivery and interest in participating in future and upcoming workshops. For instance, the VGP genome assembly pipeline on Galaxy Europe makes it intuitive for a biologist with minimal computational skills to assemble genomes and generate genome assembly assessment and visualizations of results. It is particularly relevant to students and early-career researchers in Africa, where limited computational facilities are available or accessible to non-specialists. In the future, the AfricaBP Open Institute will work with the VGP to make this assembly workflow available through the Galaxy Africa instance and other African computing platforms.

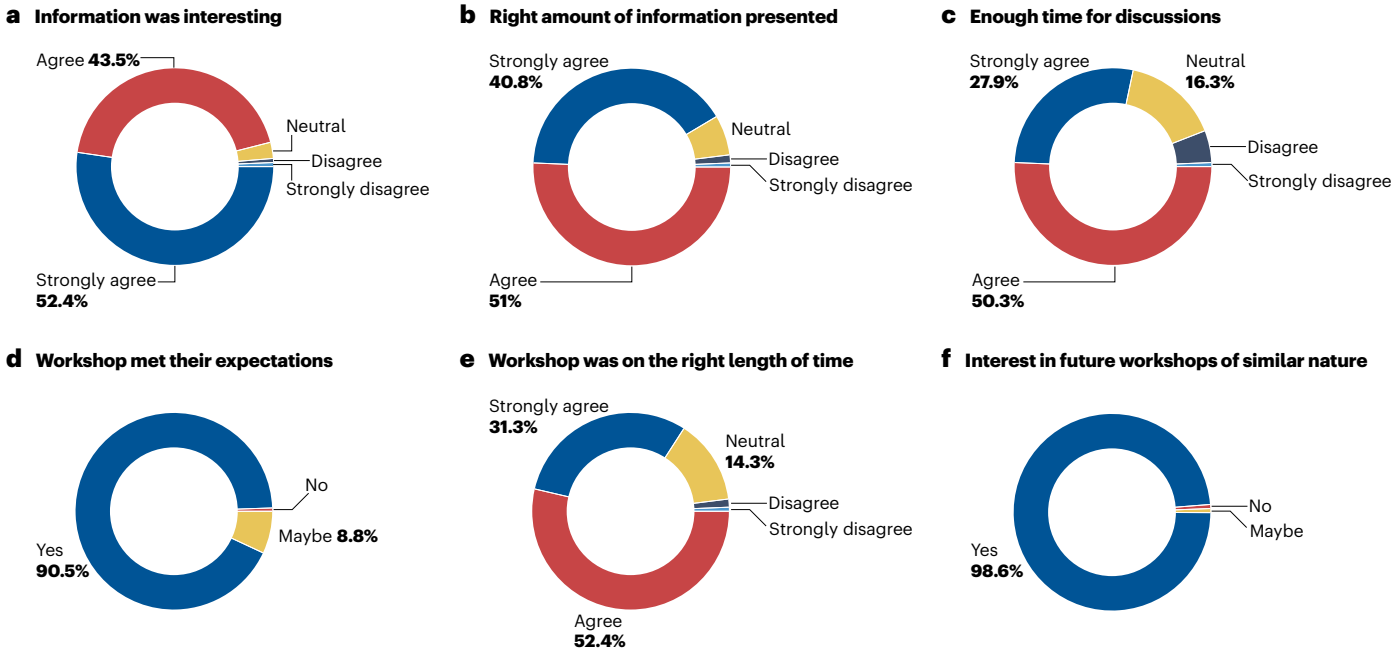

Fig. 5 | Responses of post-workshop survey to evaluate impact and outcome of the AfricaBP workshop on genomic technologies and infrastructures. Participants' responses pertaining to quality of the workshop and future interest in similar workshops. Number of respondents, 165.

It is noteworthy that the majority of the participants were from countries with active genomics research activities, such as Nigeria, Kenya and Morocco (Fig. 3). Large geographical areas of the continent, for example Angola, Namibia and Sudan, were not represented. This could be because the outreach of the AfricaBP Open Institute in these areas was not effective or simply because minimal genomics activity happens in these areas.

The AfricaBP Open Institute workshops are currently being led and delivered by African scientists. This, in turn, is helping to deliver the goal of building leadership in genomics and bioinformatics across Africa by mentoring early-career scientists, establishing key contacts and building networks. For instance, the AfricaBP has secured a platform within the African Galaxy instance and a license to host its own Research Electronic Data Capture (REDCap, https://redcap.africanbiogenome.org) as a secure survey and data capture platform. The Galaxy platform will assist scientists with limited resources or computational skills, as it is user-friendly and flexible, provides several implemented bioinformatics tools and supports hands-on workshops. Currently, the Galaxy Africa instance is supported by one of the computational clusters hosted at the Pasteur Institute of Tunis, in Tunisia. Efforts are underway within AfricaBP to secure cloud-based support from cloud providers. This will help with storage and visualization and increase processing capacity as the AfricaBP project scales.

The AfricaBP Open Institute will build on the successes of the workshops on endemic African species and biodiversity genomics technologies and infrastructure. It already has five workshops planned before the end of 2023. These include two online workshops and three regional hybrid workshops. For instance, one is an online workshop that will focus on science communication and grant writing, and it will involve a residential grant-writing exercise to apply for an identified grant. Others are regional hybrid workshops to be hosted and coordinated by the University of South Africa in South Africa, Mohammed V University in Morocco and the International Livestock Research Institute in Kenya, for the Southern, Northern, and Eastern and Central African regions, respectively. In early 2023, a regional hybrid workshop was also hosted and coordinated by the University of Port Harcourt in Nigeria. This workshop focused on the importance of biodiversity genomics in conservation of species; it included sessions tailored to emphasize the value of taxonomy in ensuring integrity of species of interest and proceeded to steps involved in the collection and processing of samples and to laboratory protocols for resolving genome sequences.

The experiences gained from the AfricaBP Open Institute biodiversity genomic technologies and infrastructure workshop, as well as the post-workshop feedback (Fig. 5), will guide future AfricaBP Open Institute workshops in bringing genomics and bioinformatics technologies closer to the African people, enabling the AfricaBP Open Institute to iteratively adapt and expand its efforts to better meet Africa's research needs. For instance, commercial genomics technology and service providers who aim to better target African biodiversity genomics researchers and organizations, at the executive and grassroots levels, could partner with the AfricaBP Open Institute to sponsor similar workshops.

Abdoallah Sharaf[1,2], Charlotte C. Ndiribe[3], Taiwo Crossby Omotoriogun[4,5], Linelle Abueg[6], Bouabid Badaoui[7,8], Fatu J. Badiane Markey[9], Girish Beedessee[10], Diaga Diouf[11], Vincent C. Duru[12], Chukwuike Ebuzome[13], Samuel C. Eziuzor[14], Yasmina Jaufeerally Fakim[15], Giulio Formenti[6], Nidhal Ghanmi[16], Fatma Zahra Guerfali[17,18], Isidore Houaga[19], Justin Eze Ideozu[20], Sally Mueni Katee[21], Slimane Khayi[22], Josiah O. Kuja[23], Emmanuel Hala Kwon-Ndung[24], Rose A. Marks[25,26], Acclaim M. Moila[27], Zahra Mungloo-Dilmohamud[28],

Sadik Muzemil[29], Helen Nigussie[30], Julian O. Osuji[31], Verena Ras[32,33], Yves H. Tchiechoua[34], Yedomon Ange Bovys Zoclanclounon[35], Krystal A. Tolley[36,37], Cathrine Ziyomo[21], Ntanganedzeni Mapholi[38], Anne W. T. Muigai[39,40] ✉, Appolinaire Djikeng[19,21,38] ✉ & ThankGod Echezona Ebenezer[41,42] ✉

[1]SequAna Core Facility, Department of Biology, University of Konstanz, Konstanz, Germany. [2]Genetic Department, Faculty of Agriculture, Ain Shams University, Cairo, Egypt. [3]Department of Cell Biology and Genetics, University of Lagos, Lagos, Nigeria. [4]Biotechnology Unit, Department of Biological Sciences, Elizade University, Ilara-Mokin, Nigeria. [5]A.P. Leventis Ornithological Research Institute, University of Jos, Jos, Nigeria. [6]Vertebrate Genome Lab, The Rockefeller University, New York, NY, USA. [7]Mohammed V University in Rabat, Rabat, Morocco. [8]African Sustainable Agriculture Research Institute, Mohammed VI Polytechnic University, Laâyoune, Morocco. [9]Rutgers University, School of Graduate Studies, Newark, NJ, USA. [10]Department of Biochemistry, University of Cambridge, Cambridge, UK. [11]Laboratoire Campus de Biotechnologies Végétales, Département de Biologie Végétale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Sénégal. [12]Department of Parasitology and Entomology, Nnamdi Azikiwe University, Awka, Nigeria. [13]Finima Nature Park, Port Harcourt, Nigeria. [14]Department of Isotope Biogeochemistry, Helmholtz Center for Environmental Research-UFZ, Leipzig, Germany. [15]Department of Agriculture, University of Mauritius, Reduit, Mauritius. [16]Bioinformatics Lab, Pasteur Institute of Tunis, Tunis, Tunisia. [17]Laboratory of Transmission, Control and Immunobiology of Infections, Pasteur Institute of Tunis, Tunis, Tunisia. [18]University of Tunis El Manar, University Campus Farhat Hached, Tunis, Tunisia. [19]Centre for Tropical Livestock Genetics and Health, Roslin Institute, University of Edinburgh, Edinburgh, UK. [20]Genomics Research Center, AbbVie, North Chicago, IL, USA. [21]International Livestock Research Institute, Nairobi, Kenya. [22]Biotechnology Research Unit, CRRA-Rabat, National Institute of Agricultural Research, Rabat, Morocco. [23]Bioinformatics Center, University of Copenhagen, Copenhagen, Denmark. [24]Department of Plant Science and Biotechnology, Federal University of Lafia, Lafia, Nigeria. [25]Department of Horticulture, Michigan State University, East Lansing, MI, USA. [26]Department of Molecular and Cell Biology, University of Cape Town, Cape Town, South Africa. [27]Inqaba Biotec, Menlo Park, Pretoria, South Africa. [28]Digital Technologies Department, University of Mauritius, Reduit, Mauritius. [29]School of Life Science, University of Warwick, Coventry, UK. [30]Department of Microbial, Cellular and Molecular Biology, Addis Ababa University, Addis Ababa, Ethiopia. [31]University of Port Harcourt, Port Harcourt, Nigeria. [32]Computational Biology Division, Department of Integrative Biomedical Sciences, IDM, CIDRI Africa Wellcome Trust Centre, University of Cape Town, Cape Town, South Africa. [33]Department of Biodiversity and Conservation Biology, University of the Western Cape, Bellville, South Africa. [34]Pan African University Institute for Basic Sciences Technology and Innovation, Nairobi, Kenya. [35]Department of Crop Sciences and Biotechnology, Jeonbuk National University, Jeonju, South Korea. [36]South African National Biodiversity Institute, Claremont, Cape Town, South Africa. [37]Centre for Ecological Genomics and Wildlife Conservation, University of Johannesburg, Johannesburg, South Africa. [38]Department of Agriculture and Animal Health, University of South Africa, Florida, South Africa. [39]Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya. [40]National Defense University–Kenya, Nakuru, Kenya. [41]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK. [42]Early Cancer Institute, Department of Oncology, School of Clinical Medicine, University of Cambridge, Cambridge, UK.

✉e-mail: awmuigai@yahoo.co.uk; appolinaire.djikeng@ed.ac.uk; thankgod1980@yahoo.co.uk

Published online: 12 September 2023

References

1. Schneegans, S., Lewis, J. & Straza, T. *UNESCO Science Report: The Race against Time for Smarter Development*, 30–77 (2021); https://unesdoc.unesco.org/ark:/48223/pf0000377433.locale=en
2. UNESCO. UIS Statistics—science, technology and innovation (accessed 15 January 2023); http://data.uis.unesco.org/
3. Agnew, A. et al. The power of data to advance the SGDs (Elsevier, 2020); https://www.elsevier.com/__data/assets/pdf_file/0004/1058179/Elsevier-SDG-Report-2020.pdf
4. Sianes, A., Vega-Muñoz, A., Tirado-Valencia, P. & Ariza-Montes, A. *PLoS One* **17**, e0265409 (2022).
5. Ebenezer, T. E. et al. *Nature* **603**, 388–392 (2022).
6. Njaci, I. et al. *Nat. Commun.* **14**, 1915 (2023).
7. Ras, V. et al. *PLoS Comput. Biol.* **17**, e1008640 (2021).
8. Moore, B. et al. *PLoS Comput. Biol.* **17**, e1009218 (2021).
9. Almarri, M. A. et al. *Cell* **184**, 4612–4625.e14 (2021).
10. El-Attar, E. A. et al. *Front. Genet.* **13**, 797465 (2022).
11. Samir Abdelhafiz, A., W L Ho, C. & Chuan Voo, T. *Wellcome Open Res.* **6**, 4 (2021).
12. Bowden, H. W., Steinhauer, K., Sanz, C. & Ullman, M. T. *Neuropsychologia* **51**, 2492–2511 (2013).
13. Tavares, N. J. *Int. J. Biling. Educ. Biling.* **8**, 319–335 (2015).
14. Woolston, C. & Osório, J. *Nature* **570**, 265–267 (2019).
15. Mahboob, A. & Lin, A. M. Y. In *Conceptual Shifts and Contextualized Practices in Education for Glocal Interaction* (eds Selvi, A. & Rudolph, N.) 197–217 (Springer, 2018).
16. Taiwo, R. O. et al. *BMC Med. Ethics* **21**, 124 (2020).
17. Nekoto, W. et al. In *Findings of the Association for Computational Linguistics (EMNLP 2020)* 2144–2160 (Association for Computational Linguistics, 2020); https://doi.org/10.18653/V1/2020.FINDINGS-EMNLP.195
18. Mlotshwa, B. C. et al. *Genet. Med.* **19**, 826–833 (2017).
19. Akindolire, M. A., Aremu, B. R. & Ateba, C. N. *Microbiol. Resour. Announc.* **8**, 01296–18 (2019).
20. da Fonseca, R. R. et al. *Mar. Genomics* **30**, 3–13 (2016).
21. Dures, S. G. et al. *Divers. Distrib.* **25**, 870–879 (2019).
22. Curry, C. J. et al. *Mol. Biol. Evol.* **38**, 48–57 (2021).
23. Lehocká, K. et al. *PLoS One* **16**, e0258714 (2021).
24. Pečnerová, P. et al. *Curr. Biol.* **31**, 1862–1871.e5 (2021).
25. Rhie, A. et al. *Genome Biol.* **21**, 245 (2020).
26. Folarin, O. A., Happi, A. N. & Happi, C. T. *Genome Biol.* **15**, 515 (2014).
27. Giovanni, M. Y. *J. Infect. Dis. Microbiol.* **1**, 1–10 (2022).
28. Gathii, J. T. *Int. Leg. Mater.* **58**, 1028–1083 (2019).
29. Pasara, M. T. & Diko, N. *Sustainability* **12**, 1419 (2020).
30. Fusacchia, I., Balié, J. & Salvatici, L. *Eur. Rev. Agric. Econ.* **49**, 237–284 (2022).
31. Helmy, M., Awad, M. & Mosa, K. A. *Appl. Transl. Genom.* **9**, 15–19 (2016).
32. Schwarze, K. et al. *Genet. Med.* **22**, 85–94 (2020).
33. Dolgin, E. *Nature* **559**, 291–293 (2018).
34. Webb, H. et al. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)* **CS19**, 1–8 (2019); https://doi.org/10.1145/3290607.3299063
35. Abdalla, M. A. & McGaw, L. J. *Front. Pharmacol.* **9**, 456 (2018).
36. Kamau, E. C. *Global Transformations in the Use of Biodiversity for Research and Development* (ed. Kamau, E. C.) 155–206 (Springer, 2022).
37. Larivière, D. et al. Scalable, accessible, and reproducible reference genome assembly and evaluation in Galaxy. Preprint at *bioRxiv* https://doi.org/10.1101/2023.06.28.546576 (2023).

Acknowledgements

The authors would like to thank N. Mulder of the University of Cape Town and M. Blaxter of the Wellcome Sanger Institute for helping to review and provide feedback on the initial draft of this manuscript.

Competing interests

A.M. is an employee of Inqaba Biotechnical Industries (Pty) Ltd. J.E.I. is an employee of AbbVie.