



Bioinformatics mentorship in a resource limited setting

Daudi Jjingo , Gerald Mboowa , Ivan Sserwadda, Robert Kakaire, Davis Kiberu, Marion Amujal, Ronald Galiwango, David Kateete, Moses Joloba and Christopher C. Whalen

Corresponding author: Daudi Jjingo, African Center of Excellence in Bioinformatics and Data Intensive Sciences, Makerere University, Kampala, Uganda. Tel.: +256-775-199-690. E-mail: daudi.jjingo@mak.ac.ug

Abstract

Background: The two recent simultaneous developments of high-throughput sequencing and increased computational power have brought bioinformatics to the forefront as an important tool for effective and efficient biomedical research. Consequently, there have been multiple approaches to developing bioinformatics skills. In resource rich environments, it has been possible to develop and implement formal fully accredited graduate degree training programs in bioinformatics. In resource limited settings with a paucity of expert bioinformaticians, infrastructure and financial resources, the task has been approached by delivering short courses on bioinformatics—lasting only a few days to a couple of weeks. Alternatively, courses are offered online, usually over a period of a few months. These approaches are limited by both the lack of sustained in-person trainer–trainee interactions, which is a key part of quality mentorships and short durations which constrain the amount of learning that can be achieved.

Methods: Here, we pioneered and tested a bioinformatics training/mentorship model that effectively uses the available expertise and computational infrastructure to deliver an in-person hands-on skills training experience. This is done through

Daudi Jjingo received his PhD in Bioinformatics from the Georgia Institute of Technology in Atlanta, USA. He is a Lecturer and Scientist in the Department of Computer Science at Makerere University and the current Director of the African Center of Excellence in Bioinformatics at Makerere University. He holds research interests in bioinformatics and data science as applied to biomedical and public health research.

Gerald Mboowa received his PhD in Human Genetics and Genomics at Makerere University in Kampala, Uganda. He is the Bioinformatics Scientist at the African Center of Excellence in Bioinformatics at Makerere University and holds research interests in Bioinformatics and Antimicrobial Resistance.

Ivan Sserwadda is currently completing his masters in bioinformatics as an NIH masters fellow at Pwani University in Nairobi, Kenya. He holds research interests in Antimicrobial resistance and bioinformatics pipelines development.

Robert Kakaire obtained his PhD in Epidemiology from the University in Athens, Georgia, USA. He is a Research Professional III in the department of Epidemiology & Biostatistics, at the Global Health Institute in the College of Public Health at the University of Georgia. He holds research interests in the community network propagation of tuberculosis.

Mariam Amujal is a Bioinformatics and Genomics PhD research fellow in the Department of Molecular Biology and Immunology, Makerere University College of Health Sciences. She holds research interest in the interaction between human genomics and HIV in Africa.

Ronald Galiwango obtained his PhD in Epidemiology and Biostatistics from the University of Georgia, Athen, USA. He is currently a postdoc scientist on the BRECA bioinformatics training program at Makerere University with research interests in Epidemic Network Modeling and modeling the interaction between air pollution and COVID-19.

David Kateete obtained his PhD in Molecular Genetics at Makerere University in Uganda. He is Senior Lecturer and Chair, Department of Immunology and Molecular Biology, Makerere University with research interests in TB lineages in Africa.

Moses Joloba obtained his PhD in Molecular Microbiology at Case Western Reserve University, Cleveland, Ohio, USA. He is a Professor and Dean School of Biomedical Sciences, Makerere University College of Health Sciences with research interests in TB molecular epidemiology and drug resistance.

Christopher C Whalen is a Distinguished Professor in Global Health within the Department of Epidemiology and Statistics at the University of Georgia, USA. He has research interests in mechanisms of the HIV-tuberculosis coinfection, treatments of HIV-associated tuberculosis, and prevention of tuberculosis as well as building capacity for research, teaching and clinical care in Africa.

Submitted: 8 June 2021; Received (in revised form): 11 August 2021

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

a few physical lecture hours each week, guided personal coursework over the rest of the week, group discussions and continuous close mentorship and assessment of trainees over a period of 1 year.

Results: This model has now completed its third iteration at Makerere University and has successfully mentored trainees, who have progressed to a variety of viable career paths.

Conclusions: One-year (intermediate) skills based in-person bioinformatics training and mentorships are viable, effective and particularly appropriate for resource limited settings.

Key words: bioinformatics; education; computational biology; mentorship; training; intermediate

Introduction

A large clinical and epidemiologic platform for research has been developed in Uganda since 1990s. Indeed, many studies have been performed that have helped shape current knowledge about HIV [1], tuberculosis (TB; [2, 3]) and malaria [4–6]. The capacity and technical expertise to perform these studies was developed quickly through viable collaborations with partners in industrialized countries and through long-term training programs. More recently, the country has become home to several National Institutes of Health (NIH) and Wellcome Trust funded projects geared towards enhanced understanding of Human Health and Heredity in Africa (H3Africa; [7]). These include the Collaborative African Genomics Network (CAfGEN [8, 9]), the Integrated Biorepository of H3Africa Uganda (IBRH3AU [8]), the H3BioNet Uganda node [10], the Nurturing Genomics and Bioinformatics Research Capacity in Africa (BReCA) and programs exploring the ethical and social landscape of genomic research in Uganda and Africa [11]. However, the capacity by African researchers to fully leverage all the data from those projects through basic science research has lagged behind because of insufficient infrastructure, skills and the high cost of experimental basic science research.

Bioinformatics, which we operationally define as a discipline concerned with the acquisition, storage, analysis, interpretation and dissemination of biological data, most often molecular sequences like DNA, RNA and amino acids—sits at the interface between clinical research and basic science as it draws upon concepts of population science and the knowledge and theory of biological and computational science. It thus offers an opportunity to bridge these fields using computational, statistical, artificial intelligence and biological techniques that are feasible in low and middle income countries (LMIC). With bioinformatics, there is now a new field open to researchers in LMIC where infrastructure needs may be more readily met. The digital nature of bioinformatics offers an opportunity to reduce the cost of basic science research. It is thus particularly important and appropriate for low-resourced settings like Uganda in their quest to build their research capacity and discovery potential within the life sciences. This possibility has been rendered more feasible by the recent advent of high-throughput sequencing and computing technologies [12].

Sufficiently skilled bioinformaticians are still few and far between within LMICs, but they are critical for the completion of the scientific promise of bioinformatics. There have been several approaches in trying to fill this critical gap [13]. The most common is organizing short term trainings like workshops or seminars, usually lasting only a few days or at most a couple of weeks [13–17]. This model is common mostly because trainers from Highly Industrialized Countries or resource-rich environments can only afford limited amounts of time away from their home institutions. The greatest limitation to this

model of training, however, is that bioinformatics covers such a wide domain [18, 19] that it can hardly be sufficiently taught and properly assimilated by trainees in a few days or weeks [17–19]. The second approach has been online tutorials that usually last a few weeks or months [20, 21]. In the African context, the most successful of these has been H3ABioNet's Introduction to Bioinformatics course (IBT; [20]). It has, however, been established that learning is best delivered in longer term in-person settings as is the case for high schools [17, 22–26], undergraduate [27, 28] and graduate degree programs [29]. This model is constrained by limited resources like expert trainers, computing infrastructure and effective accreditation bureaucracies. Such resources are essential to initiate and sustain fully accredited bioinformatics degree education programs, which also require years of commitment. Indeed, this is evidenced by the existence of only few such programs in Africa [29]. Although the need for bioinformatics competencies in Africa has been well established [10], the best approaches for delivering this content to trainees to build a sufficient workforce of bioinformaticians in resource limited settings are yet to be fully worked out.

We pioneered an intermediate training approach that tries to optimize the use of the available expertise, limited computational infrastructure and time constraints of both trainees and trainers in resource-limited environments. It ensures a sustained period of skills based, hands-on and in-person instruction with close mentorship that spans a year. It provides a path to both applied bioinformatics practice in research and industry as well as a smooth transition to graduate education in bioinformatics.

Program organization and structure

To optimize resources, the program leveraged the already existing academic and organizational infrastructure at the two partnering Universities (Figures 1 and 2): The University of Georgia (UGA) in the United States and Makerere University in Uganda. The two universities are jointly studying the transmission dynamics of HIV and TB in African urban settings using modern methods of molecular epidemiology, which include bioinformatics. As a fundamental principle of collaboration shared by both institutions, the Program Directors (PDs) of the research project sought to develop a parallel bioinformatics training and mentorship program, with only minimal additional costs.

In this structure, the PDs provide overall scientific direction and strategic planning for the program. As for implementation, they are responsible for securing funding for the mentorship program that supports local program mentors, the development of a competency-based progressive curriculum, in-person tutorials and materials and equipment for homework and group

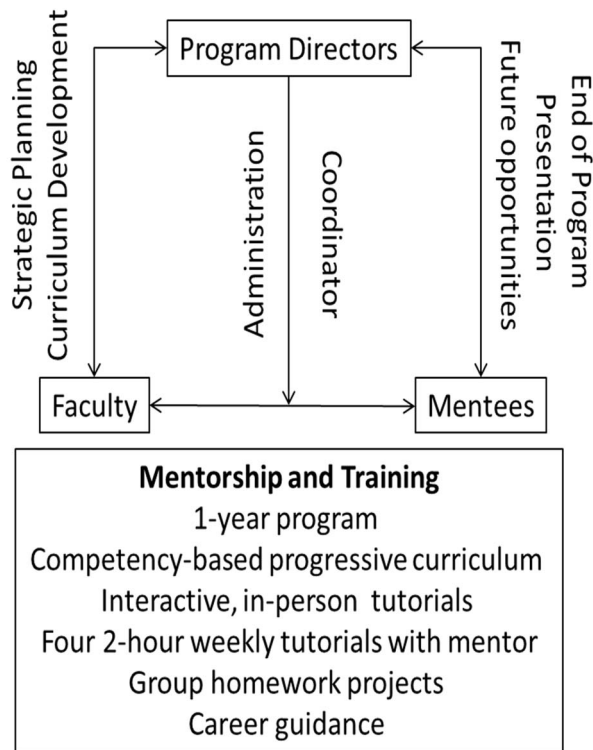


Figure 1. Program organogram, illustrating the three major program components and how they relate to each other.

projects. The fiscal aspects of the project were managed through the College of Public Health at the University of Georgia and the secretariat of the Uganda Society of Health Sciences (USHS) at Makerere University.

At the implementation level, the local program mentors include Senior Faculty and Assistant Instructors. Senior Faculty are responsible for curriculum formulation, periodic curriculum review, recruiting the mentees and providing oversight for the instruction, mentorship and assessment activities. They also reserve institutional support and resources such as classroom space and access to computational infrastructure for the program in addition to identifying the Assistant Instructors. In turn, the Assistant Instructors are responsible for helping with classroom instruction, providing one-on-one skills-based mentorship and assessment of trainees. Mentors are selected from local faculty at Makerere University with prior training in bioinformatics and from existing programs with bioinformatics related projects such as CAFGEN and H3Africa. Typically, the senior faculty were PhD graduates who had obtained mentorship skills as part of their experiences on the PhD programs. Others had been mentors in other mentorship programs such as the MUII Mentorship program of MRC-UVRI [30–33]. Among other factors, mentors are motivated by having mentees eventually help with their ongoing research comprising bioinformatics. It also helps them meet some of their teaching requirements if they are PhD students and provides a practiced path towards potential faculty positions for postdocs. Furthermore, faculty are interested in exploring locally appropriate models of training that could eventually fill some of the skills gap in the region. At the beginning of the program, instructors were given train-the-trainer courses to bring them up to speed with the course materials. In addition, over the course of

the program, instructors are encouraged to attend seminars and workshops in bioinformatics or take advanced courses online as a way of remaining updated about new developments in the dynamic field. Because all the above described individuals already exist within the university, as are the classroom and computational facilities, the mentorship is able to be executed at only a minimum incremental cost.

Program delivery modalities

The Mentorship program was designed to optimize flexibility while ensuring the highest standards of instruction and covering all the critical competencies of bioinformatics in the African context (Table 1; [28, 34]). These competencies are those important for research themes of immediate relevance to the African biomedical landscape, which is dominated by infectious diseases like HIV, TB and malaria. They include genome analysis and annotation (for viral and bacterial bioinformatics), molecular epidemiology of pathogens (for drug resistance surveillance and detection), sequence analysis and phylogenetics (for HIV and TB evolution) and research ethics [11].

The various bioinformatics domains are encapsulated into a progressive curriculum of self-contained modules that are delivered in logical sequence (Figure 3). The curriculum was designed to teach both the scientific knowledge underlying bioinformatics and a practical set of skills that are essential in applying that knowledge. Basic topics are taught first before advanced topics. Course delivery involves both theory and practical hands-on sessions that are handled by both the local Senior Faculty and Assistant Instructors using a total of 8 class/contact hours each week for 49 weeks. Outside of those timetabled hours, the Assistant Instructors offer extended practical sessions and office hours as and when needed by the mentees. Mentees start each mentorship cycle together and are given routine reading and practical assignments by which they are assessed at the end of each module. Performance was measured through session attendance, engagement with instructors and peers, as well as performance on practical assignments. Graduation was held and certificates of completion were awarded to only those who attended at least 70% of the mentorship sessions and successfully completed at least 70% of the assignments. The program seeks timely feedback at the end of each module to ensure quick and flexible adjustment of subsequent modules for improved delivery during rather after each mentorship cycle. On top of the canonical bioinformatics domains (defined as common cores at benchmarked graduate bioinformatics programs—bioinformatics programming, primary sequence databases and sequence analysis), the mentorship also includes journal club sessions to give mentees competence in literature surveys, academic writing and grant writing. These are included because of their importance for the mentees' career growth. At the end of the mentorship year, overall feedback is collected from the graduating mentees and discussed between the instructors, senior faculty and PDs. This ensures continuous improvements in program management, curriculum content and delivery styles between mentorship cycles. Trainees are awarded certificates at the end and are helped in finding subsequent placements on full time masters or PhD programs, research laboratories or industry (Table 2).

By running the program over 1 year, we were able to pace the delivery of instructional material and gauge student understanding along the way. This approach allowed for an adaptive style that could be tailored for each individual student. The dynamic and flexible scheduling of the 8 instructional hours each week was carefully implemented to ensure flexibility for

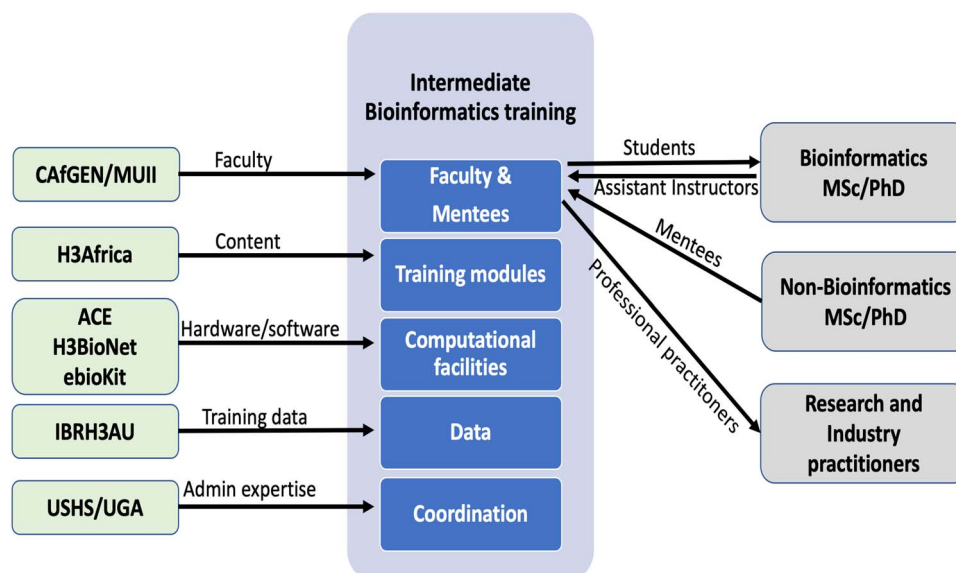


Figure 2. Diagram illustrating how existing organizational infrastructure (green boxes) is leveraged for program success (blue boxes) and how the program relates to MSc/PhD programs and professional/industry research practice (grey boxes).

Table 1. The specifically selected bioinformatics domains of the mentorship, covering all critical bioinformatics competencies. It is delivered in 4 quarters of 3 months each

Quarter	Course	Major concepts
First quarter	Molecular biology and DNA sequencing (platforms and chemistry) Command line computing and scripting	Central dogma, DNA structure and function, early DNA sequencing methods, high-throughput sequencing (HTS) methods Linux operating system, Unix filesystem and navigation, Unix commands and syntax, shell scripting
Second quarter	Genome assembly and research ethics Genome analysis and genome annotation	De novo and reference-based Genome assembly, reads quality metrics, genome assembly tools, primary sequence databases, ethics of genomics Workflows for microbial and human genome analysis; read alignment, variant calling and variant annotation
Third quarter	Sequence analysis and phylogenetics Literature review and presentation Scientific writing	Sequence alignment, phylogenetic tools and phylogenetic interpretation Journal clubs; paper reading, paper presentation and paper discussion Manuscript writing; paper arrangement, content and purpose of paper sections, drawing paper figures and graphics.
Fourth quarter	Bioinformatics programming and R statistical computing	R studio, R data input/output, R base, management and use of R packages and BioConductor

both instructors and mentees since the mentorship program is meant to fit within their other mainstream commitments. In this way, the 1-year program proves more feasible in logistics and schedule for students who may have limited time because they are in another degree program or professionals with work commitments or who do not have the resources to take a full masters or doctoral program in bioinformatics. That intermediacy, which nonetheless covers all important competencies [34], fills a critical training niche that is even more important in resource limited settings.

The program has leveraged infrastructure and resources from other projects within the University's eco-system. These resources include trainees from the CAFGEN who have been co-opted as instructors [8, 9], the ebioKit bioinformatics server [35], the biobanking experience offered by the IBRH3AU, supplemental online instruction from H3ABioNet's Introduction to Bioinformatics course (IBT; [20]) and high performance computing access provided by the African Center of Excellence

(ACE) in Bioinformatics and Data Intensive Sciences [36]. Furthermore, trainees get opportunities to rotate within laboratories in the departments of Immunology & Molecular Biology, Medical Microbiology and the ACE [36, 37]. The former hosts a MiSeq next-generation sequencing platform among other equipment important in genomics and bacterial culture capabilities, whereas the later provides the computational bioinformatics and visualization infrastructure. The three facilities offer the students the opportunity to learn different aspects of experimental microbial and molecular biology while also applying genomics and bioinformatics in their exploration and analysis.

On the program, the mentee numbers kept growing in each succeeding mentorship cycle, but was a maximum of eight in the latest. This suggests a maximum capacity of about 10 with the currently available pool of instructors. Expansion of the program is possible by co-opting some of the graduated mentees and some current trainees of the

Bioinformatics Competences	1 st Quarter	2 nd Quarter	3 rd Quarter	4 th Quarter
Biology and Molecular Biology	✓		✓	
Biological data generation technologies	✓			
Command line Computing and Scripting	✓	✓		✓
Bioinformatics tools and software		✓	✓	✓
Online Bioinformatics and Sequence databases		✓	✓	
Ethics of Genomics		✓		
Scientific Communication and writing			✓	
Team Work	✓	✓	✓	✓
Statistical Research Methods				✓
Bioinformatics Programming	✓			✓
Continuing Professional Development				✓

Figure 3. Order and quarterly distribution of bioinformatics competencies training.

MSc/PhD programs in bioinformatics at Makerere University as instructors.

Student selection and performance

Special attention was paid to selection of the mentees. Each prospective trainee submitted an application that was reviewed by the program faculty. In selecting trainees, the faculty developed two main criteria that were considered critical for successful mentorship: (i) rigorous academic background in the biological or quantitative sciences and (ii) focus and commitment of prospective students to learning bioinformatics. Eligible applicants can be individuals that completed their undergraduate, individuals doing their Masters, PhD or postdoc in fields other than bioinformatics who are interested in picking up some bioinformatics skills or practicing professionals or researchers. Being nested within a large university, the program was able to recruit students from across a multi-collegial pool. Their backgrounds included both life sciences (human medicine, public health, veterinary medicine and agriculture) and physical sciences (physics, mathematics, statistics and computing). That diversity in background not only captured all the building blocks of bioinformatics, but also maximized cross pollination and synergistic learning between the trainees. This happened because

team-based assignments, discussions, journal clubs and interactive lecturers allow mentees to share information and skills from their varied backgrounds. Training students from an array of university colleges and research establishments also sets up the diffusion of bioinformatics into those diverse environments where it can be deployed to a variety of applications. The same effect resulted in diffusing bioinformatics into existing research programs such as CAfGEN [8], the Makerere University-Uganda Virus Research Institute (UVRI) Centre of Excellence for Infection & Immunity Research and Training (MUII) [30–33], UVRI and BReCA where some of the alumni of the mentorship end up.

As a way to assess the success of the program, the directors tracked the subsequent training or research positions obtained by the alumni of the program. Moreover, program mentors took an active role in securing positions for their trainees and for documenting the post-tutorial success of the mentees. For example, a number of the program alumni won scholarships on the NIH sponsored masters in bioinformatics program (BReCA) and another got a full scholarship for her PhD on the same program. Another was admitted to a fulltime bioinformatics masters degree program at the Birkbeck University of London, the United Kingdom. Three contributed excellent, award-winning talks to the USHS annual conference that aims to showcase current research and learning among Uganda's health sciences community. The additional exposure to writing and

Table 2. Trainee inventory and post-tutorial positions

Trainee	Mentorship year	Background	Current status
Trainee 1	2017	MSc in Immunology & Clinical Microbiology	PhD in Bioinformatics (Baylor College of Medicine, United States and Makerere University, Uganda)
Trainee 2	2018	BSc Biotechnology	MSc in Immunology & Clinical Microbiology (Makerere University, Uganda)
Trainee 3	2018	BSc Biochemistry	MSc in Bioinformatics (Makerere University, Uganda)
Trainee 4	2018	BSc Biology	PhD in Bioinformatics (Makerere University, Uganda) & Lectureship at Busitema Public University
Trainee 5	2018	BSc Medical Laboratory Science	MSc in Bioinformatics (Pwani University, Kenya)
Trainee 6	2018	BSc Medical Lab Science	MSc in Immunology & Clinical Microbiology (Makerere University, Uganda)
Trainee 7	2018	MSc Molecular Biology	PhD in Molecular Biology (Makerere University, Uganda)
Trainee 8	2018	BSc Biotechnology	MSc in Bioinformatics (Makerere University, Uganda)
Trainee 9	2019	BSc Biomedical lab	MSc in Bioinformatics (Birkbeck University of London, the UK)
Trainee 10	2019	BSc Medical Lab Science	MSc in Bioinformatics (Makerere University, Uganda)
Trainee 11	2019	BSc Biochemistry	MSc in Bioinformatics (Makerere University, Uganda)
Trainee 12	2019	BSc Biomedical Sciences	MSc in Bioinformatics (Makerere University, Uganda)
Trainee 13	2019	BSc Biomedical lab	MSc in Bioinformatics (Makerere University, Uganda)
Trainee 14	2019	BSc Biochemistry	MSc in Immunology & Clinical Microbiology (Makerere University, Uganda)
Trainee 15	2019	BSc Biochemistry	MSc in Bioinformatics (Makerere University, Uganda)
Trainee 16	2019	BSc Biomedical lab	MSc in Bioinformatics (Makerere University, Uganda)

scientific communication has enabled some of mentees to compete and receive travel awards for both oral and poster presentations at different regional genomics and bioinformatics conferences such as the H3Africa and ISCB/ASCB [8].

The program has created a pool of trainees ready to feed into different graduate programs (like Master's and PhD in bioinformatics) or advance careers in research, industry and instructorships on the same or different mentorship programs (Table 2). Indeed, several of the trainees have secured competitive scholarships to pursue their education in Kenya, Uganda, Europe and the United States. Similarly, the instructors have had the opportunity to improve their pedagogical and mentorship skills in a novel context.

Discussion and conclusion

We established a 1-year intermediate tutorial/mentorship program in bioinformatics that sits between short term workshops (few days to a few months) and full degree (3–4 years) programs. It covers all important bioinformatics competencies [34] and fills a critical bioinformatics training niche that is especially important in resource limited settings.

There is no doubt that the duration of training generally correlates with the amount and depth of knowledge and skills acquired by trainees. Indeed, this is why degree programs lasting 3–4 years are necessary and a preferred path for many young people seeking to join most fields. However, this path is not open for the most motivated and capable people in LMICs because most prospective students lack adequate resources for tuition covering 3 or 4 years. To accommodate this group of motivated people, universities and research institutions have organized and offered short courses in bioinformatics, lasting days to weeks. Although these short-courses and workshops play a meaningful role for individuals seeking an introduction to bioinformatics, they fall short in building the sustainable capacity for research.

At either of those two ends along this spectrum of training, educational institutions in LMIC face challenges or encounter limitations because bioinformatics training requires scientific expertise and computational infrastructure. To support degree programs, universities need a faculty with specific training in bioinformatics who sustain an active research program that serves as a basis for education and training. Since bioinformatics is a new field, many established researchers do not have the expertise to mentor and direct student education in bioinformatics. Furthermore, bioinformatics is multidisciplinary so it is a challenge to find an academic home for these scientists.

Between the full degree program and short-course approach, there is an important niche for intermediate training whose resource requirements are more manageable in LMICs. To create this intermediate training, we developed a tutorial program that was built upon close and continuous mentoring for 1 year. Moreover, the tutorials were organized around practical skills in bioinformatics. This pragmatic approach streamlined the theoretical content that was taught but helped to spur on further interest in bioinformatics. As the trainees developed analytic and quantitative skills, they became more useful to their research mentors. Moreover, this kind of training and mentorship made trainees more competitive for graduate training scholarships and fellowships, thereby mitigating against the funding challenges experienced by these young individuals.

We found that leveraging existing resources in academic and research structures within the university can help support and grow an intermediate 1-year in-person mentorship program in bioinformatics. Indeed, the benefits worked both ways. As the program tapped into the expertise of available faculty, this faculty was the direct beneficiary of trainees as they applied their newly developed skillset to faculty research.

Admittedly, such a program does have its own challenges: as one might expect, the requirement for close in-person mentorship limits the mentor-mentee ratio, which means scaling the program can be a slow process given the limited number of potential instructors with the appropriate level of bioinformatics

skills and knowledge. So far, attempts to address this have included co-option of program graduates as instructors. Second, in locations where computational infrastructure might be in acute shortage, alternative platforms like cloud computing or remote access to regional server resources like ACE [36, 37] would have to be sought. Alternatively required computation within assignments would have to be toned down to what is doable on just a laptop or desktop computer. Third, although there are criteria for trainee assessment, there is still room for improvement. For example, future assessments will be further strengthened with formulation of a rubric based on the bioinformatics competencies and a corresponding score scale. Furthermore, the program has not yet had the opportunity to be tested at other institutions/Universities. Depending on the maturity of their computational, faculty and parallel research projects environments, they might have to make adjustments to attain a format that best suits their context. Finally, as with all other educational programs, this mentorship was challenged by the emergence of COVID-19, which meant the transition of several of the didactic sessions to online platforms. Nevertheless the small class-size the mentorship sessions (10 maximum) left room for some socially distanced in-person engagements, which mitigated against the complete elimination of the valuable in-person mentorship.

The continuity of such a program is critical for this type of training to have a positive effect. We believe that the current program that was pioneered at Makerere University has laid a strong foundation for the future. It is guaranteed to have a continuous supply of eager new students since potential recruits can come from diverse science backgrounds—physical, technological and life science backgrounds. Those recruits can come from different levels—bachelors, masters, PhD degrees, post-docs, researchers and those in industry as long as their objective is to attain functional capacity in bioinformatics. The training can be undertaken alongside other formal non-bioinformatics graduate degree programs and other commitments like professional research or industry practice. Since the top performers in one class can be co-opted as tutors in future training cycles, there is a renewable supply of potential competent faculty and assistant instructors. The cost of the program can be kept to a minimum by leveraging existing infrastructure and expertise located in a University environment. This minimal cost is more affordable to trainees as tuition, which is critical for the program's financial sustainability. Furthermore, the program design included options for both self-sponsorship and endowed fellowships. Presence of both types of mentees on the program suggests in-house ability for the program to generate some resources for its sustainability. This model of training and mentorship is thus an efficient and effective approach for delivering this exceptional skill to a broad group of persons. As faculty become more available and the challenges discussed here continue to be addressed, the program has the potential to grow to 20 or so mentees in each mentorship cycle. In addition, it will hopefully eventually start at other locations/universities once they have made minimal adjustments to suit their specific resource contexts.

Key Points

- We developed and tested an intermediate bioinformatics training/mentorship model that lies between short courses on the one hand and fulltime degree courses on the other.

- It comprises a few in-person lecture hours each week, guided personal coursework over the rest of the week, group discussions and continuous close mentorship and assessment of trainees over a period of 1 year.
- This model has now completed its third iteration at Makerere University, in Uganda, and has successfully mentored trainees, who have progressed to a variety of viable career paths (Table 2).
- One-year (intermediate) skills based in-person bioinformatics training and mentorships are viable, effective and particularly appropriate for resource limited settings.

Data availability statement

All data used for the manuscript is included within the tables in the manuscript.

Acknowledgements

The African Center of Excellence in Bioinformatics (ACE) and data sciences at Makerere University, which provided some of the requisite computational resources. The Collaborative African Genomics Network (CAfGEN), which trained some of the program faculty. The Infectious Diseases Institute (IDI), which supported some of the program faculty. The Center for Computational Biology at Uganda Christian University supported some of the program faculty.

Funding

This work was supported by the National Institutes of Health, Fogarty International Center [grant numbers: 1D43TW010045-01, 1U2RTW010672-01].

References

1. Serwadda D, Mugerwa RD, Sewankambo NK, et al. Slim disease: a new disease in Uganda and its association with HTLV-III infection. *Lancet Lond Engl* 1985;2:849–52.
2. Srikanthiah P, Lin R, Walusimbi M, et al. Elevated HIV seroprevalence and risk behavior among Ugandan TB suspects: implications for HIV testing and prevention. *Int J Tuberc Lung Dis* 2007;11:168–74.
3. Whalen CC, Zalwango S, Chiunda A, et al. Secondary attack rate of tuberculosis in urban households in Kampala, Uganda. *PLoS One* 2011;6(2):e16137.
4. Achan J, Kakuru A, Ikilezi G, et al. Antiretroviral agents and prevention of malaria in HIV-infected Ugandan children. *N Engl J Med* 2012;367:2110–8.
5. Byakika-Kibwika P, Lamorde M, Okaba-Kayom V, et al. Lopinavir/ritonavir significantly influences pharmacokinetic exposure of artemether/lumefantrine in HIV-infected Ugandan adults. *J Antimicrob Chemother* 2012;67:1217–23.
6. Dick GWA, Kitchen SF, Haddow AJ. Zika virus (II). Pathogenicity and physical properties. *Trans R Soc Trop Med Hyg* 1952;46(5):521–34.
7. Mulder N, Abimiku A, Adebamowo SN, et al. H3Africa: current perspectives. *Pharmacogenomics Pers Med* 2018;11:59.
8. Mboowa G, Mwesigwa S, Katagirya E, et al. The collaborative African genomics network (CAfGEN): applying genomic

- technologies to probe host factors important to the progression of HIV and HIV-tuberculosis infection in sub-Saharan Africa. *AAS Open Res* 2018;1:3.
9. Mlotshwa BC, Mwesigwa S, Mboowa G, et al. The collaborative African genomics network training program: a trainee perspective on training the next generation of African scientists. *Genet Med* 2017;19:826–33.
 10. Aron S, Gurwitz K, Panji S, et al. H3ABioNet: developing sustainable bioinformatics capacity in Africa. *EMBnetjournal* 2017;23:886.
 11. Ali J, Cohn B, Mwaka E, et al. A scoping review of genetics and genomics research ethics policies and guidelines for Africa. *BMC Med Ethics* 2021;22:39.
 12. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. *Mol Cell* 2015;58:586–97.
 13. 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.
 14. Ahmed A, Awadallah AA, Elmahdi MT, et al. Blended bioinformatics training in resource-limited settings: a case study of challenges and opportunities for implementation. *bioRxiv* 2018;431361.
 15. Cummings MP, Temple GG. Broader incorporation of bioinformatics in education: opportunities and challenges. *Brief Bioinform* 2010;11:537–43.
 16. Tastan Bishop Ö, Adebisi EF, Alzohairy AM, et al. Bioinformatics education—perspectives and challenges out of Africa. *Brief Bioinform* 2015;16:355–64.
 17. Via A, Rivas JDL, Attwood TK, et al. Ten simple rules for developing a short bioinformatics training course. *PLoS Comput Biol* 2011;7:e1002245.
 18. Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? A proposed definition and overview of the field. *Methods Inf Med* 2001;40:346–58.
 19. Schneider MV, Watson J, Attwood T, et al. Bioinformatics training: a review of challenges, actions and support requirements. *Brief Bioinform* 2010;11:544–51.
 20. Gurwitz KT, Aron S, Panji S, et al. Designing a course model for distance-based online bioinformatics training in Africa: the H3ABioNet experience. *PLoS Comput Biol* 2017;13(10):e1005715.
 21. 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.
 22. Louisa Wood and Philipp Gebhardt Bioinformatics goes to school—new avenues for teaching contemporary. *Biology* 2013;9(6):e1003089.
 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:e1002244.
 24. Kovarik DN, Patterson DG, Cohen C, et al. Bioinformatics education in high school: implications for promoting science, technology, engineering, and mathematics careers. *CBE Life Sci Educ* 2013;12:441–59.
 25. Machluf Y, Yarden A. Integrating bioinformatics into senior high school: design principles and implications. *Brief Bioinform* 2013;14:648–60.
 26. Wefer SH, Sheppard K. Bioinformatics in high school biology curricula: a study of state science standards. *CBE Life Sci Educ* 2008;7:155–62.
 27. Madlung A. Assessing an effective undergraduate module teaching applied bioinformatics to biology students. *PLoS Comput Biol* 2018;14(1):e1005872.
 28. Sayres MAW, Hauser C, Sierk M, et al. Bioinformatics core competencies for undergraduate life sciences education. *PLOS ONE* 2018;13:e0196878.
 29. Machanick P, Tastan BÖ. How to establish a bioinformatics postgraduate degree programme—a case study from South Africa. *Brief Bioinform* 2015;16:346–54.
 30. Cose S, Bagaya B, Nerima B, et al. Immunology in Africa. *Trop Med Int Health* 2015;20:1771–7.
 31. Elliott A, Nerima B, Bagaya B, et al. Capacity for science in sub-Saharan Africa. *The Lancet* 2015;385:2435–7.
 32. Akello M, Coutinho S, Mboowa MG, et al. Experience of introducing monitoring research activities to promote quality research among MUII-Plus trainees. *AAS Open Res* 2019;3:57.
 33. Biraro IA, Driciru E, Drago CK, et al. MUII-PLUS mentorship programme: lessons from the baseline survey. *AAS Open Res* 2019;3:37.
 34. 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:e1005772.
 35. Hernández-de-Diego R, de Villiers EP, Klingström T, et al. The eBioKit, a stand-alone educational platform for bioinformatics. *PLoS Comput Biol* 2017;13.
 36. 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.
 37. ACE – African Center of Excellence in Bioinformatics & Data Sciences <https://ace.idi.co.ug/>