

The collaborative African genomics network training program: a trainee perspective on training the next generation of African scientists

Busisiwe C. Mlotshwa, MSc¹, Savannah Mwesigwa, MSc², Gerald Mboowa, MSc², Lesedi Williams, MSc¹, Gaone Retshabile, MSc¹, Adeodata Kekitiinwa, MBChB, MMed³, Misaki Wayengera, MBChB, MSc², Samuel Kyobe, MBChB, MMed², Chester W. Brown, MD, PhD^{4,5,9}, Neil A. Hanchard, MBBS, PhD^{4,5,6}, Graeme Mardon, PhD^{4,7}, Moses Joloba, MBChB, PhD², Gabriel Anabwani, MBChB, MMed⁸ and Sununguko W. Mpoloka, PhD¹

Purpose: The Collaborative African Genomics Network (CAfGEN) aims to establish sustainable genomics research programs in Botswana and Uganda through long-term training of PhD students from these countries at Baylor College of Medicine. Here, we present an overview of the CAfGEN PhD training program alongside trainees' perspectives on their involvement.

Background: Historically, collaborations between high-income countries (HICs) and low- and middle-income countries (LMICs), or North–South collaborations, have been criticized for the lack of a mutually beneficial distribution of resources and research findings, often undermining LMICs. CAfGEN plans to address this imbalance in the genomics field through a program of technology and expertise transfer to the participating LMICs.

Methods: An overview of the training program is presented. Trainees from the CAfGEN project summarized their experiences, looking specifically at the training model, benefits of the program, challenges encountered relating to the cultural transition, and program outcomes after the first 2 years.

Conclusion: Collaborative training programs like CAfGEN will not only help establish sustainable long-term research initiatives in LMICs but also foster stronger North–South and South–South networks. The CAfGEN model offers a framework for the development of training programs aimed at genomics education for those for whom genomics is not their “first language.”

Genet Med advance online publication 6 April 2017

Key Words: African scientists; bioinformatics; genomics research; North–South gap; training model

Increasingly, the global pattern of disease burden is shifting from infectious to chronic.^{1,2} In 2013, 6 of the 10 diseases listed by the US Centers for Disease Control and Prevention as leading causes of death in the United States were chronic diseases.³ Current research⁴ focuses on identifying the genetic determinants of disease with the understanding that identification of these causative factors may help improve diagnostic indicators and also provide potential therapeutic targets for future pharmaceutical intervention.⁵

The African continent, which bears a disproportionate burden of disease, has an evident deficit in the advanced genetic and genomics technologies that can lead to the development of novel approaches to disease control. The paucity of well-trained, knowledgeable African genomics scientists is one factor hindering the development of African research in this field.⁶ A 2011 report⁷ linking research output to local researchers

in-country found that, with the exceptions of South Africa, Senegal, and Nigeria, the numbers were disappointingly low, averaging less than 2,000 researchers per country for the majority of sub-Saharan countries.

Africa has the fastest growing population in the world.⁸ Coupled with a high incidence of communicable diseases, such as malaria, tuberculosis, and HIV, and an increase in noncommunicable diseases, such as diabetes, cancer, stroke and heart disease, it has become even more critical that the appropriate technical, technological, and infrastructural expertise be developed to address these challenges.

The global North–South divide is based on socioeconomic development, with the North comprising high-income countries (HICs) from North America and Europe and the South comprising low- and middle-income countries (LMICs) from Africa, Asia, and South America. Historically, North–South

¹Department of Biological Sciences, Faculty of Sciences, University of Botswana, Gaborone, Botswana; ²Department of Medical Microbiology, College of Health Sciences, Makerere University, Kampala, Uganda; ³Baylor College of Medicine Children's Foundation–Uganda, Kampala, Uganda; ⁴Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA; ⁵Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA; ⁶ARS/USDA/Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; ⁷Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX 77030; ⁸Botswana-Baylor Children's Clinical Centre of Excellence, Gaborone; ⁹Current affiliation: Genetics Division, Department of Pediatrics, University of Tennessee Health Science Center, Le Bonheur Children's Hospital, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. Correspondence: Sununguko W. Mpoloka (mpoloka@mopipi.ub.bw)

collaborations have been criticized for the inequitable distribution of resources and research findings.^{9–11} Scarce research funding and inadequate training opportunities have been cited as some of the challenges affecting this imbalance.^{7,12} Nevertheless, collaborative training programs between HICs and LMICs have been successfully implemented; one such example is a partnership between Makerere University and the Karolinska Institutet in Sweden, which produced 44 PhD graduates and offshoot collaborations.¹³

In light of this inequity, the Human Heredity and Health in Africa (H3Africa) initiative¹⁴ was established as a partnership between the US National Institutes of Health (NIH), the UK-based Wellcome Trust, and the African Society of Human Genetics (AfSHG).¹² This 5-year, \$74 million initiative aims to narrow the gap in genomics expertise and research on the African continent by developing the capacity for health-related genomics research because previous African participation has been minimal.¹⁵ Biomedical research in Africa has traditionally been limited to specimen collection, with subsequent specimen shipments to developed countries where virtually all of the major investigations were conducted: the “postal research”¹⁶ or “helicopter science”^{17,18} models.^{17,18} Developing the necessary genomics research proficiencies will allow African scientists to set appropriate frameworks to conduct comprehensive genomics studies addressing health issues relevant to African populations and to compete effectively on the global platform.¹⁵ The relevance of genomics to research in both chronic and infectious diseases makes it an essential inclusion in the path to establishing African health.¹⁹

Under the aegis of the H3Africa initiative, the Collaborative African Genomics Network (CAfGEN) was created with the specific aim of understanding the underlying host genomic factors influencing progression of HIV and HIV–tuberculosis infection in pediatric populations in Botswana and Uganda while contributing to the infrastructure, expertise, and sustainability of genomics capacity within the H3Africa consortium.^{20,21}

CAfGEN comprises five sites (**Figure 1**): two African pediatric clinics, two African universities, and a college in the United States. The clinics—Botswana–Baylor Children’s Clinical Center of Excellence and the Baylor College of Medicine Children’s Foundation–Uganda—provide clinical expertise for patient recruitment and sample collection. The universities—Makerere University and the University of Botswana—furnish the program with PhD trainees, process patient samples, and provide the infrastructure and opportunities to sustain genomics research endeavors through the establishment of graduate programs in genetics and genomics, in addition to sequencing and bioinformatics facilities in the home institutions. Baylor College of Medicine (BCM) in Houston, Texas, affords trainees access to genomics expertise, including hands-on experiences in the laboratory, didactic coursework, data analysis, and a variety of other educational activities.

Beyond the scientific objectives, CAfGEN seeks to develop and nurture young African scientists in the field of genomics, thereby increasing research capacity and sustainability in Africa through a program of technology and expertise transfer. This will empower them to investigate diseases relevant to their communities. As part of that mandate, CAfGEN has established a long-term training program in which African PhD trainees undertake 2 years of didactic training and experiential learning at BCM in genomics and bioinformatics methodologies. On completion of training at BCM, trainees return to their respective home institutions to fulfill the requirements for a PhD. As trainees progress through the program, they participate in experiences that include H3Africa consortium meetings, annual BCM departmental genetics/genomics retreats, and a variety of seminars relevant to genomics and bioinformatics. Trainees are introduced to a dynamic community of scholars, with whom they establish mentoring and peer relationships during their time at BCM, which will continue throughout their careers. Continuous program assessments by trainees and faculty are used to improve and refine the CAfGEN program

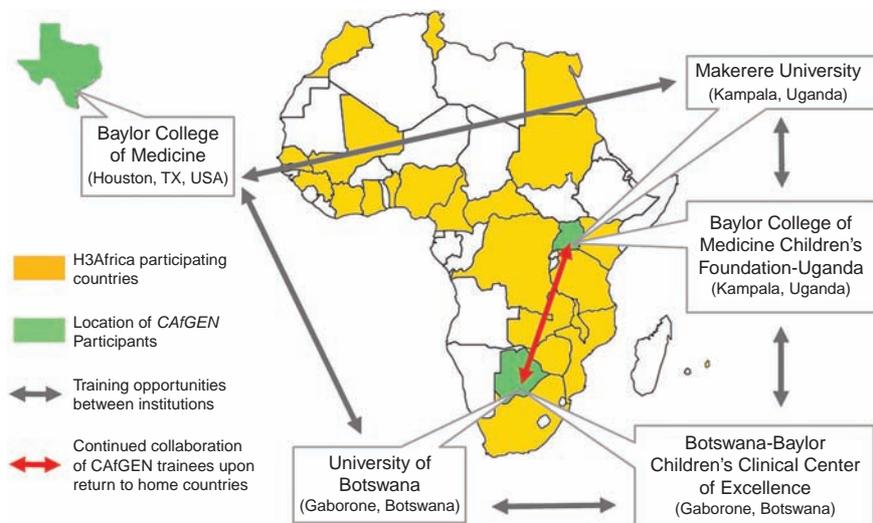


Figure 1 Levels of interaction among the five member institutions of CAfGEN. Adapted with permission from <http://www.worldatlas.com/webimage/countrys/africa/afoutl.htm> and <http://www.worldatlas.com/webimage/countrys/namerica/usstates/outline/tx.htm>. Copyright 2015 by Valnet Inc.

and to ensure that it realizes its objectives. The expectation is that CAfGEN trainees will be instrumental in raising awareness of the importance of genomics and bioinformatics as research tools within their home institutions and communities, and will help develop and expand educational tracks in these disciplines as part of the groundwork towards creating a self-sustaining genomics research culture in Africa.

The longer-term CAfGEN training goals include establishing and enhancing undergraduate, graduate, and faculty education in genetics/genomics and providing opportunities for long- and short-term training of scientists and technicians from African universities. Additional goals are establishing genetic/genomic technologies and supporting laboratory and physical infrastructure for large-scale analyses of common diseases in Africa. This training program ensures retention of African genomic scientists on the continent, with the aim of building a sustainable, critical number of researchers who will influence African universities and governments to invest in, support, and improve biomedical research infrastructure as a means of finding diagnostic and therapeutic solutions to Africa-relevant health and disease.²²

In this article, six trainees share their collective training experiences from their time at BCM. Trainees consider the current outcomes of the program, as well as outline the challenges anticipated in conducting genomics research in their respective African countries. The CAfGEN training program is presented as a potential translatable model for future training programs in human genetics and/or genomics, in and beyond Africa. Trainees appraise the method of training, the applicability of the acquired skills in their respective home environments, the benefits of the program to trainees and their communities, and

make recommendations to improve this model for future training programs.

PROGRAM DESCRIPTION

Trainee selection and demographics

The six current trainees, three each from Botswana and Uganda, were recruited through the two universities via both internal and external advertising. Selection criteria included a postgraduate degree, with preference given to those in the fields of genetics, cell biology, biostatistics, and computer science. Candidates who were short-listed based on their academic background (**Table 1**) and an essay describing their motivation, as well as testimonials provided by three academic referees, were subsequently invited to oral interviews conducted by a panel of university staff. Potential scholarship recipients identified at this stage were required to successfully register with the university as PhD students and submit a research proposal based on the CAfGEN research objectives. Scholarship recipients were also required to pass BCM admission standards before the application process was considered complete. The selection process was carried out over several months at Makerere University and the University of Botswana, hence the arrival of trainees at BCM at three different time points. All six trainees have continued through the program; the first five have reached the point of returning to their respective home institutions and the sixth has a year remaining at BCM.

Curricular components

Genomics is characterized by “big data,” and its analysis and the coursework that CAfGEN trainees undertake is geared toward

Table 1 Academic qualifications and work experience of trainees

Trainee	Academic qualifications	Work experience	Prior bioinformatics experience (none, basic, intermediate, advanced) ^a
01	BSc (biological sciences), MSc (quantitative genetics and genome analysis)	Laboratory manager (4 years) Laboratory coordinator (1 year) Laboratory research assistant (3 years)	Basic
02	MB, ChB (bachelor of medicine and bachelor of surgery), MMed (medical microbiology)	Junior house officer (1 year) Head of infection control (1 year) Medical officer (2 years)	Basic
03	BSc (biomedical science), MSc (medical microbiology, immunology, and molecular biology)	Clinical laboratory coordinator (1 year) Research scholar (1 year) Laboratory research assistant (3 years)	Basic
04	BSc (biomedical laboratory technology), MSc (immunology and clinical microbiology)	Assistant lecturer (3 years) Laboratory supervisor (7 years)	Basic
05	BSc (biological sciences), MSc (molecular and cellular biology)	Lecturer (1 year) Laboratory demonstrator (2 years) Teaching assistant (3 years)	Basic
06	BSc (biochemistry), MSc (molecular biology and biotechnology)	Assistant lecturer (4 years) Research associate (5 years)	Basic

^aBasic: familiarity with bioinformatics tools such as BLAST and public bioinformatics Web-based resources (e.g., NCBI, ENSEMBL); intermediate: familiarity with basic coding or scripting in any programming language; advanced: large-scale analysis of biological data and proficiency in any programming language.

Table 2 CAfGEN trainee program coursework and seminars

Coursework		
Course title	Term ^a	Major concepts/tools
Genetics A; Genetics B	1 and 2	Mendelian genetics, statistical methods, population genetics, model systems, genome analysis
Application to Biology of Computation	2	Epigenomics, network biology, sequence analysis
Practical Introduction to Programming for Scientists	3	BioPython, NumPy, Matplotlib, GUI programming, networking
Gene Regulation	3	Transcription in bacteria, chromatin, pre-messenger RNA splicing, translational regulation mechanisms
Introduction to Biostatistics	4	Association, power, nonparametric methods, survival analysis
Human Genetics	4	Pedigrees, genomic disorders, mitochondria, complex inheritance, population genetics
Computer-Aided Discovery Methods	4	Cancer biology, ChIP-seq data analysis using MACS2, Integrative Genomics Viewer
Bioinformatics and Genome Analysis	4	ChIP-Seq data analysis using seqMINER and MEME-ChIP, DAVID, UCSC Genome Browser
Ethics for Professionals ^b	All	Human-subjects research, scientific misconduct, care and use of animals, plagiarism
Seminars		
Series	Term ^a	Content
Exome Sign-out	All	DNA diagnostics, interpretation of whole-exome sequencing data in the evaluation of Mendelian disorders
Molecular and Human Genetics Seminars	All	Weekly seminar series featuring external speakers presenting current genetics/genomics-related research

^aDuration of a term is 6 weeks. ^bNational Institutes of Health prerequisite for federal grant funding.

Table 3 Laboratory rotations in year 1 of the CAfGEN program

Rotation	Objectives	Tools/methods
DNA and RNA Extraction (Hanchard Laboratory—USDA/ARS Children’s Nutrition Research Center)	Designing genetics studies. How to conduct quality DNA and RNA extraction. Understanding SNP array genotyping, Sanger sequencing and NGS techniques. QC of data.	QIACube RNA extraction. DNA quantification with qPCR and spectrophotometry. Gel electrophoresis. Sequencher, PLINK, GenomeStudio.
WES and RNA-Sequencing (Chen Laboratory)	Understanding NGS. Molecular diagnosis of disease through targeted whole-exome sequencing. Setting up an NGS variant analysis pipeline. Using public databases as a source of controls in genetic studies.	OMIM, RetNet, ExAC, 1000 Genomes, dbSNP, CHARGE. Mutation Taster, SIFT, POLYPHEN. Bash scripting.
Sanger Sequencing and WES (Human Genome Sequencing Center [HGSC])	Understanding next-generation sequencing. Understanding sample quality control procedures. Extraction of nucleic acid from blood plasma. Whole-exome sequencing. Setting up bioinformatics pipelines for NGS QC and analysis. NGS variant discovery best practice. Compute resource utilization.	qPCR nucleic acid quality control. Gel electrophoresis. WES with the Illumina platform (MiSeq and HiSeq). Bash scripting. BWA-MEM. SAMTools. GATK.
GWAS and SNP Genotyping (Leal Laboratory)	Designing genetic studies. How to call variants and evaluate their functionality in GWAS and NGS studies. Performing analysis of population-based data for complex traits as well as rare variant data.	GWAS QC, Association and Complex trait analysis using PLINK, CASSI, BEAM3, GenABEL, SEQPower.
Population Genetics and WES (Fuli Laboratory—Computational and Integrative Biomedical Research Center)	Understanding bioinformatics workflows for NGS. Understanding GATK best practices. Understand various variant callers used in NGS. Running QC metrics on NGS data in comparison with public data (1000 Genomes Project). Understanding how to annotate variants.	Bash and python scripting to extract information from VCF files. Compute resource management with PBS scripting. NGS sequence analysis and QC with BWA-MEM, GATK, SAMTools, VCFTools.
eQTL, SNPs, RNA Sequencing (Hanchard Laboratory—USDA/ARS Children’s Nutrition Research Center)	Understanding differences between RNA-seq and DNA-seq alignment. Running differential expression analysis on RNA-seq data.	Sequence alignment with TopHat. Differential expression analysis with Cumberbund, Bowtie2, Cufflinks, R (EdgeR and DESeq).

eQTL, expression quantitative trait loci; GATK, Genome Analysis Toolkit; GWAS, genome-wide association studies; NGS, next-generation sequencing; QC, quality control; SNP, single-nucleotide polymorphism; VCF, variant call format; WES, whole-exome sequencing.

equipping trainees with the skills that will enable them to conduct genomic studies.

The first year of training consists primarily of didactic coursework (Table 2) and laboratory rotations (Table 3). Trainees also attend seminars and workshops focused on a variety of genetic and/or genomics topics (Table 2). The aim of many of these lectures is to expose trainees to available informatics tools used in genomics research and their application to human subject

research. Many of the tools that trainees are exposed to are freely available, open-access software, which they can continue to use as they start their careers at their home institutions. Newly acquired programming and computational skills provide a strong foundation that the trainees will continue to build on in order to enhance their skills to a level from which they can independently develop their own analysis tools as needed (Table 2). The program design is such that the laboratory

rotations intentionally overlap in scope to reinforce learning. At the end of the rotations, each trainee will have received instruction in designing genetic studies and the theoretical and practical aspects of next-generation sequencing from sample intake to production of analysis-ready variants. They will also have received basic training in the bioinformatics requirements for running next-generation sequencing, variant discovery, quality control, and annotation. The second year concentrates on analysis of data from the CAfGEN cohorts; during this period, trainees are placed in laboratories that complement their research interests. A timeline of the sequence of events for each trainee from arrival at BCM (**Figure 2**) emphasizes that the rotations and coursework at the core of the program can be tailored to each intake or trainee, depending on research interests; non-core program laboratories are included in the program at trainees' requests.

Trainees deliver presentations on relevant topics in monthly genomics journal clubs and during their laboratory rotations; this promotes critical reading of the literature and practice in interpreting scientific results, both of which are essential skills for developing scientists. Trainees also have the opportunity to interact with prominent visiting researchers,^{23,24} who graciously share their own training experiences. Visiting researchers have included prominent members of the Independent Expert Committee of H3Africa who have research interests

focusing on next-generation genomics approaches and population genomics. These researchers are always well published and provide inspiration to the trainees regarding the potential impact they can make in their own areas of research. Monthly conferences, in which clinical genetics cases that have been solved by whole-exome sequencing are presented, give trainees a window into the victories and shortcomings of genomics as applied to clinical medicine.

In their second year of training, trainees have already begun to apply the various proficiencies acquired during their first year in a holistic manner by analyzing sequencing data sets generated from CAfGEN cohorts in Africa. The trainees have attended H3Africa consortium meetings, competing successfully within the pool of H3Africa trainees for travel scholarships to present their work at these esteemed events. Trainees also participate monthly via teleconference in the H3Africa Fellows' Club Webinar Series, a platform for trainees that enhances genomics networking, training, and research in Africa.

Mentorship

Mentorship at BCM is available at two levels: from the CAfGEN principal investigators (PIs) and through the laboratory rotations. CAfGEN PIs have weekly meetings with the trainees to discuss their research projects, laboratory rotations, and coursework, and they ensure that the trainees are on track to meet the

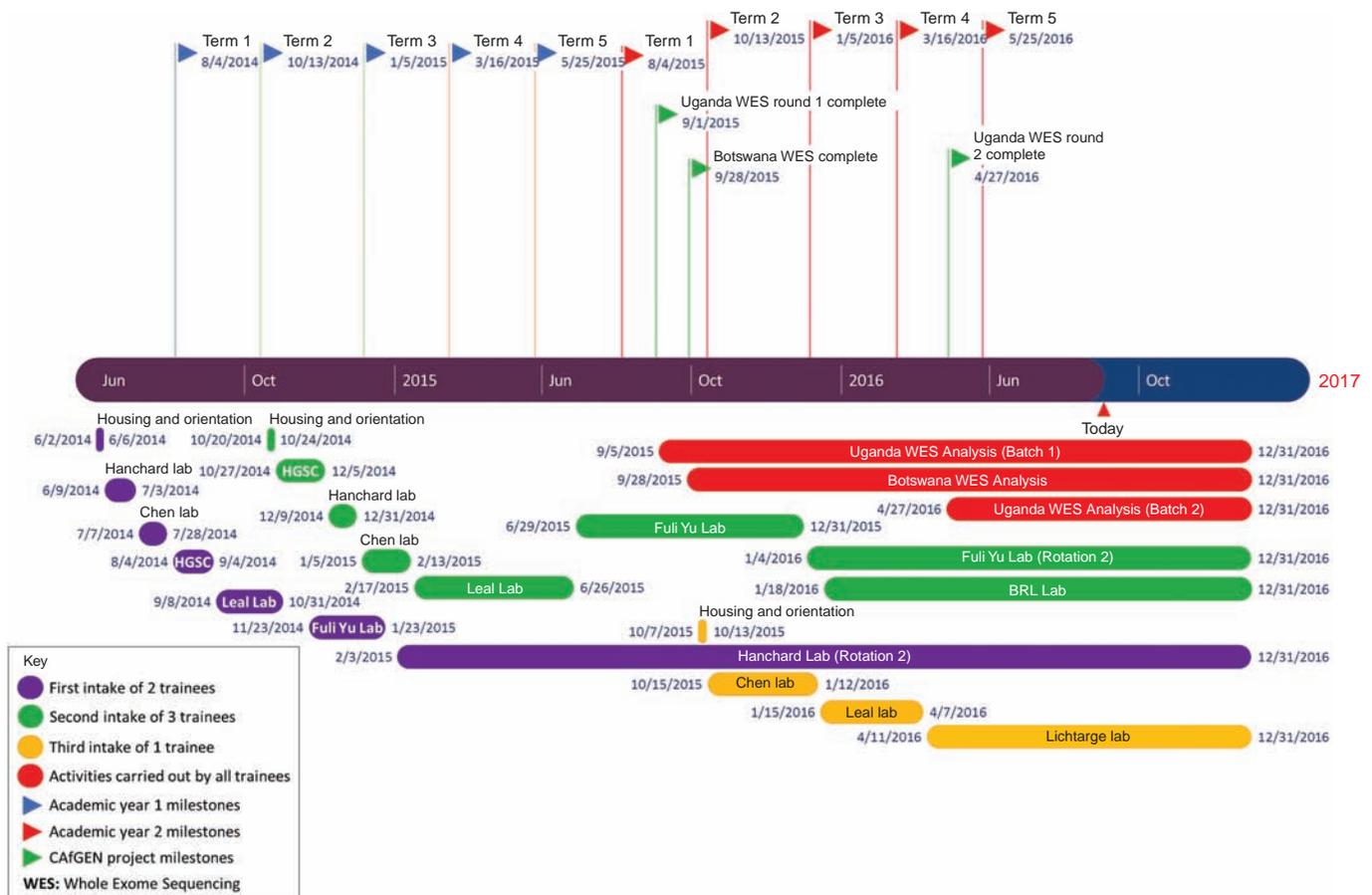


Figure 2 CAfGEN activities over the 2-year training period.

degree requirements of their home institutions. Trainees also receive mentoring from the postdoctoral fellows and laboratory PIs they interact with throughout their various laboratory rotations. Dependent on the particulars of each trainee's research project, this mentoring is intended to extend beyond the time spent in that actual laboratory to the period when trainees return home. Trainees also have the opportunity to source mentors with similar research interests or career paths from the greater H3Africa/NIH pool during the biannual H3Africa consortium meetings. This offers trainees a comprehensive mentorship network that covers technology, career development, work-life integration, and work processes.²⁵

TRAINEE PERSPECTIVES

Trainees' perspectives were noted during monthly focus groups in which laboratory rotations, coursework, and general well-being were discussed with project mentors. The following descriptions represent a consensus of the trainees' views.

Benefits of the program

Formal collaborations can result from personal relationships between individuals.²² In that respect, the long-term training program implemented by CAfGEN is more beneficial than short-term training programs or workshops. Trainees are more likely to form stronger bonds, both among themselves and with their mentors at BCM, because of the longer training period. This is an added advantage for continued mentoring after the trainees return home. Interactions with postdoctoral fellows and graduate students from the training institution permit trainees to become part of a network composed not only of mentors but also of early and mid-career scientists in genomics, enabling present and future collaborations between institutions. Another benefit of the program is the bringing together of trainees from different countries; the 2-year training period fosters strong working relationships that can potentially lead to lifelong research partnerships between African institutions. The duration of the program also enables trainees to take advantage of available graduate courses, seminars, and workshops that are not available at their home institutions, as shown in **Table 2**.

Sociocultural considerations

Trainees find their attention divided between family commitments back home, obligations to the program, and meeting home institution requirements. Finding the appropriate balance among the three, especially with the distance entailed, has been a true test and an exercise in compartmentalization. Social isolation in academic life has been a challenge; coursework for CAfGEN reflects a reduced selection of the mandatory classes in the Molecular and Human Genetics program at BCM, and formation of quality relationships between trainees and their US counterparts has been hindered by limited opportunities for interaction between the two groups. The wide array of options and complexities involved in settling on appropriate service providers for necessities such as electricity and rental insurance, among others, has also required some adjustment.

Negotiating the differences between their home countries and the United States in health-care insurance has been a taxing undertaking. Homesickness has been an issue for all trainees, especially those from Botswana, for whom a mid-program visit home has not been built into their stay. Assimilating to life in Houston has led to some interesting experiences—not least among them an American Thanksgiving and “Black Friday.” Trainees have found Houston to be a multicultural city, and its students come from all over the world. The ethos of the city is an interesting hodge-podge of practicality and extravagance, with high-rise buildings under construction everywhere, a booming international restaurant culture, and offerings of events in the visual arts and music, both paid and free, on an almost-daily basis.

Challenges

One of the greatest challenges of the program is the time constraints of the various rotations versus the competencies that trainees are expected to acquire during these periods. There are also differences in teaching styles between African institutions and graduate-level BCM courses, resulting in a training experience with a steep learning curve. The utility of large data sets to identify genetic causes of disease has resulted in a plethora of tools for their analysis, with additional software becoming available as technology strives to keep up with advances in the medical field. Accordingly, it is possible that some of the skills instilled by the program will require adaptation in the not-very-distant future.

Trainees' home institutions lack a computational/bioinformatics infrastructure, attributable largely to limited funding and lack of expertise; CAfGEN and H3Africa are changing the latter and the H3Africa Bioinformatics Network (H3ABioNet; <http://h3abionet.org>) is addressing the former by developing a core bioinformatics infrastructure, so-called “nodes,” in 15 African countries and providing bioinformatics training in Africa at the graduate level and in the form of various internships and courses related to computing education. Returning trainees can address the remaining gap by lobbying for domestic funding and investment. Few institutions in Africa have departments of genetics and bioinformatics, largely because of the previously discussed shortcomings. Trainees will need to introduce these courses and programs aimed at stimulating enthusiasm among undergraduate students in their institutions, further ensuring sustainability.

Upon returning to home institutions trainees will take on the twin challenges of establishing genomic research (especially at the comparatively young University of Botswana) and enhancing comprehension of the prerequisites and benefits of genomics in their respective communities, although CAfGEN makes considerable efforts that involve community outreach and education. Library services as extensive as those of the Texas Medical Center are not common in Africa, and access to current issues of high-impact journals may be an important topic to address with home institutions. Further challenges include Africa's relatively few opportunities for postdoctoral training

in the genomics field to support advanced career development and establishing functional research groups that are locally funded in trainees' respective countries.

DISCUSSION

Lessons learned

Although the program is designed for 2 years, all trainees agreed that additional time at BCM would be beneficial to allow trainees sufficient time to become proficient with the various bioinformatics tools, identify their interests, and further develop proposals. Moreover, because bioinformatics analyses require basic knowledge of the UNIX command line interface, trainees felt that familiarity with the command line prior to undertaking laboratory rotations and coursework should be considered a prerequisite.

Sharing progress reports with various course directors initially posed a few challenges; all trainees endorsed the use of cloud server storage as helpful in sharing trainee progress reports and the course directors' reviews. Keeping itemized trainee journals to document progress helps to update the home universities on training progress and will serve as important reference material in the future. The cloud server is also useful for sharing trainee resources, such as links to online courses and relevant papers. Although Skype calls are used to make presentations to home institutions, there are some connectivity issues and the differences in time zones make it difficult to coordinate between institutions.

Certain difficulties should not be overlooked when designing an intricate training program such as CAfGEN, primarily, financial logistics of the transition to the United States and establishment of permanent living arrangements in advance. Protocols for transferring trainee stipends abroad should be established before trainees depart from their home institutions because bureaucratic procedures sometimes complicate and delay transfer of funds. Costs such as travel emergency funds should also be considered in the event that trainees need to travel home for unforeseen reasons. Establishing relationships with organizations that offer temporary accommodation to incoming international students, such as Houston Homestay (<http://www.homestayhouston.com>), could prove a solution to facilitating this transition. Host families can help new students find accommodations, open bank accounts, and familiarize themselves with the public transportation system and culture. Providing trainees with extra funds for "settling in" (e.g., deposits for apartments, connecting electricity, and acquiring basic furniture) could also go a long way in helping them establish themselves.

Trainee reintegration at their home universities

Upon completion of the first round of the CAfGEN training program, six African scientists will have been trained in genomics. To promote retention of trainees upon completion of the program, trainees will be assimilated into faculty positions in their institutions. During this period, they will transition through positions such as research fellows and later

into tenured faculty positions (e.g., lecturers) according to the research pathways stipulated by their home institutions. Both Makerere University and the University of Botswana have progressed positively toward establishing these positions, and genomics undergraduate and graduate programs are currently being designed in consultation with the BCM faculty. To ensure proper reintegration, trainees will be provided with protected time for research following their return. Several other avenues, such as re-entry grants, are being investigated to encourage the trainees' focus on the field and on the new knowledge gained. CAfGEN is currently working with H3ABioNet to ensure the creation of bioinformatics resources in both Botswana and Uganda so that trainees can continue their research upon their return to Africa. Similarly, the laboratory infrastructure is being improved in both countries, with equipment enabling Sanger and next-generation sequencing acquired and in use at the University of Botswana and Makerere University, respectively.

Future implications

CAfGEN is not a dual-institution, degree-granting program; trainees are awarded degrees by their home institutions in Africa. However, taking into consideration the long-term goals of the program and the global health initiatives at BCM, this is a direction that CAfGEN should consider, especially as the program gains momentum with subsequent groups of trainees. The model used by CAfGEN emphasizes a highly condensed 2-year training period exclusively undertaken at BCM that utilizes state-of-the-art facilities and expert faculty. The networks envisaged by CAfGEN will ensure collaborative research and continued training opportunities, and sharing information and experiences will improve the understanding of the genetic architecture of the African population with the goal of promoting human health. A United Nations report in 2014 predicted that "The future of humanity is increasingly African," noting that by the end of this century approximately 40% of all humans (and nearly half of all children) will be African.²⁶⁻²⁸ Health and environment remain important aspects of this rapidly increasing population, and genomics promises to improve our understanding of the health and diseases relevant to the African population.

Ultimately, the growth of genomics capacity in Africa is dependent on training African scientists who will commit to returning to Africa to develop sustainable genomics research. CAfGEN's model of genomics training will enable African scientists to apply bioinformatics and genomics in the pursuit of solutions to health challenges faced by their diverse and rapidly growing population.

ACKNOWLEDGMENTS

CAfGEN is funded by NIH grant 1U54AI110398. We appreciate the contribution of the following CAfGEN institutions: Baylor College of Medicine, Houston, Texas; Baylor Children's Clinical Centers of Excellence in Uganda and Botswana; Makerere University, Kampala; and University of Botswana, Gaborone. We are also grateful to the H3Africa Consortium.

DISCLOSURE

Dr. Wayengera reports other from Restrizymes Biotherapeutics (U) LTD, outside the submitted work. The other authors declare no conflict of interest.

REFERENCES

1. US Department of State. *Infectious and Chronic Disease*. <http://www.state.gov/eoes/intlhealthbiodefense/id/>. Accessed 28 January 2016.
2. Khoury MJ, Mensah GA. Genomics and the prevention and control of common chronic diseases: emerging priorities for public health action. *Prev Chronic Dis*. 2005. http://www.cdc.gov/pcd/issues/2005/apr/05_0011.htm. Accessed 28 January 2016.
3. Centers for Disease Control and Prevention, National Center for Health Statistics. *Leading Causes of Death*. <http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm>. Accessed 28 January 2016.
4. Chong JX, Buckingham KJ, Jhangiani SN, et al.; Centers for Mendelian Genomics. The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *Am J Hum Genet* 2015;97:199–215.
5. Ginsburg GS, Willard HF. Genomic and personalized medicine: foundations and applications. *Transl Res* 2009;154:277–287.
6. Folarin OA, Happi AN, Happi CT. Empowering African genomics for infectious disease control. *Genome Biol* 2014;15:515.
7. Irikefe V, Vaidyanathan G, Nordling L, Twahirwa A, Nakkazi E, Monastersky R. Science in Africa: the view from the front line. *Nature* 2011;474:556–559.
8. Kaneda T, Bietsch K. 2015 *World Population Data Sheet*. Population Reference Bureau: Washington, DC, 2015. Accessed 27 January 2016.
9. Edejer TT. North-South research partnerships: the ethics of carrying out research in developing countries. *BMJ* 1999;319:438–441.
10. Chetwood JD, Ladep NG, Taylor-Robinson SD. Research partnerships between high and low-income countries: are international partnerships always a good thing? *BMC Medical Ethics* 2015;16:36.
11. Nordling L. Research: Africa's fight for equality. *Nature* 2015;521:24–25.
12. 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.
13. Sewankambo N, Tumwine JK, Tomson G, et al. Enabling dynamic partnerships through joint degrees between low- and high-income countries for capacity development in global health research: experience from the Karolinska Institutet/Makerere University Partnership. *PLoS Med* 2015;12:1–8.
14. H3Africa. *Harnessing Genomic Technologies Toward Improving Health in Africa: Opportunities and Challenges*. 2011. http://h3africa.org/images/PDF/h3africa_whitepaper.pdf. Accessed 27 January 2016.
15. H3Africa Consortium, Rotimi C, Abayomi A, et al. Research capacity. Enabling the genomic revolution in Africa. *Science* 2014;344:1346–1348.
16. Costello A, Zumla A. Moving to research partnerships in developing countries. *BMJ* 2000;321:827–829.
17. Chu KM, Jayaraman S, Kyamanywa P, Ntakiyiruta G. Building Research Capacity in Africa: Equity and Global Health Collaborations. *PLoS Med* 2014;11:1–4.
18. Kaplan M. Genomics in Africa: avoiding past pitfalls. *Cell* 2011;147:11–13.
19. Siwo GH, Williams SM, Moore JH. The future of genomic medicine education in Africa. *Genome Med* 2015;7:47.
20. US Department of Health and Human Services. NIH Research Portfolio Online Reporting Tools (RePORT). Project information: Collaborative African Genomics Network (CAFGEN). https://projectreporter.nih.gov/project_info_description.cfm?aid=8990819&icde=29793947&ddparam=&ddvalue=&ddsub=&cr=4&csb=default&cs=ASC. Accessed 03 June 2016.
21. H3Africa. Collaborative African Genomics Network (CAFGEN). 2013. <http://h3africa.org/consortium/projects/16-projects/89-cafgen>. Accessed 3 June 2016.
22. Coloma J, Harris E. From construction workers to architects: developing scientific research capacity in low-income countries. *PLoS Biol* 2009;7:1–4.
23. Awadalla P. High-coverage sequencing of real populations reveals the underlying processes shaping genomes and disease. Molecular and Human Genetics Seminar Series, Baylor College of Medicine, Houston, TX, 3 February 2015.
24. Bustamante CD. Opportunities and challenges for medical and population genomics in the personal genome era. Molecular and Human Genetics Seminar Series, Baylor College of Medicine, Houston, TX, 12 April 2016.
25. Farren, C. *Eight Types of Mentors: Which Ones Do You Need?* 2006. http://www.masteryworks.com/newsite/clientimpact/%5Cdownloads%5CArticle3_EightTypesofMentors-WhichOnesdoYouNeed.pdf. Accessed 08 September 2016.
26. UNICEF. Annual Report 2014. http://www.unicef.org/publications/index_82455.html. Accessed 28 January 2016.
27. Ashford, Lori S. *Africa's Youthful Population: Risk or Opportunity?* 2007; <http://www.prb.org/Publications/Reports/2007/AfricasYouthfulPopulation.aspx>. Accessed 28 January 2016.
28. Orchard CA, Curran V, Kabene S. Creating a culture for interdisciplinary collaborative professional practice. *Med Educ Online* 2005;10:1–13.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

© The Author(s) (2016)