Chapter 9

The H3Africa Consortium: Publication Outputs of a Pan-African Genomics Collaboration (2013 to 2020)

Ananyo Choudhury, Dhriti Sengupta, Shaun Aron, and Michèle Ramsay

1 Introduction

Developing a complex collaborative program across multiple African countries, characterized by a diversity of government models, laws, cultures and norms, resources and skills, climatic conditions and health challenges is a daunting task. In 2010, discussions began to work toward developing a model to support a pan-African genomics initiative to ensure that Africa was not left behind in the post-genomics era with the possibility of precision medicine on the horizon. By examining the opportunities and challenges, a white paper was drafted titled: Harnessing Genomic Technologies Toward Improving Health in Africa: Recommendations for the Human Heredity and Health in Africa (H3Africa) Initiative (https://h3africa.org/index.php/about/white-paper). This was presented to the leadership of the Wellcome Trust and the National Institutes of Health, resulting in committed funding and calls for funding applications.

The H3Africa Consortium had its first meeting in August 2012 in Addis Ababa with just 7 supported projects. As of March 2020, the H3Africa Consortium has grown to over 50 funded projects (1 central coordinating center; 14 collaborative centers and 16 research projects; 2 ethics collaborative centers and 8 ethics projects; 4 pilot biorepositories and 1 bioinformatics collaborative center and 4 bioinformatics training grants), with over 500 consortium members and the equivalent of 170M USD in investment. Figure 9.1 shows the geographic location of the different projects on the continent. The African Academy of Science has become a key partner taking on specific governance roles, including managing access to the resources developed by the Consortium.

It is a challenging task to assess and quantify the impact and outputs of such a complex and multi-faceted Consortium. These include data generation, discovery and new knowledge generation, resource development (research and ethics guidelines, infrastructure, capacity strengthening through skills development and biobanks with biospecimens, development of bioinformatics tools and pipelines), information for health policy development and targeted interventions, patents and peer-reviewed publications. Importantly, the
H3Africa Consortium contributes large African genome datasets from populations across the continent to explore insights into demographic histories of migration and admixture across Africa. These data are available to the global scientific community for further analysis.
In this chapter, we have limited ourselves to assessing the publications oeuvre of the H3Africa Consortium (January 2013 to February 2020) under specific themes that reflect the areas of substantive activity. Each paper was assessed and the acknowledgements and funding attribution scrutinized to identify publications by H3Africa-associated researchers. We divided the papers into three broad categories: i) H3Africa Core Publications, including the marker paper and H3Africa workshop meeting reports, ii) Perspectives, editorials and reviews about the H3Africa Consortium more generally, and iii) the extended reach of H3Africa. The core publications are cited in the reference list, whereas the complete list of PubMed IDs is provided in Figure 9.2. The authors take full responsibility for the potential misclassification of publications, but have attempted to be objective and fair. We apologize in advance for any omissions or distress caused in this regard.

2 Identification of Publications and Categorization

We performed a PubMed database (https://www.ncbi.nlm.nih.gov/pubmed) search for publications potentially originating out of the H3Africa Consortium. We used “H3Africa” as a search term along with a list of NIH grant IDs associated with NIH funded H3Africa projects. The grant IDs searched were - HG006938, DK116913, AI110398, HG007479, HG006947, HG007480, HG009826, HL141011, HG009824, MH096754, Al136677, MH15484, HG009784, HD094658, HG007044, MH115485, HG008226, HG008222, HG008224, HG009810, HG009790, HG007008, HG007438, HG007051, HG006941, HG009780, TW010677, TW010672, TW010673, TW010679, HG006939, AI110398, HG007480, AI110422, HG007465, AI110466, HG007472, HG007459, HG007628, HG007654, HG007092, HG006941, GO10273, GO10275 and HG009822. The search was conducted on the 20th of February 2020. In addition to the 371 unique publications identified by the search, two additional key papers (in press at the time of writing) were reviewed (Figure 9.3). Based on a manual scrutiny of the research subject matter, funding and acknowledgements these 373 publications were categorized into the following groups:

H3Africa Core Publications: Publications that are the direct outcome of original research from an H3Africa funded project, with acknowledgement of the specific grant. In this category, the papers have been divided into eight different subject-related categories as follows: marker papers and cohort descriptions; governance, ethics, community engagement and biobanking; epidemiological studies on diseases, traits and risk factors; disease-associated behaviour and awareness studies; genetics and genomics studies; bioinformatics and
FIGURE 9.2 PMIDs of 371 publications associated with H3Africa

genomics capacity development; and microbiome and pathogen studies (Figure 9.4).

H3Africa Perspectives, Editorials and Reviews: Publications written by H3Afria members about the H3Africa Consortium and its role in enhancing genomics research in Africa.
Extended Reach H3Africa Publications: Publications with one or more author(s) who are fully or partly supported by H3Africa funds, whose contributing research for the paper may be partly funded by H3Africa, and who have contributed to large multi-authored papers not directly related to a project that was funded under the H3Africa Consortium, and acknowledging funding under the H3Africa Consortium umbrella.
Allocation of 169 core H3Africa publications to different subject areas (January 2013 to 20 February 2020)

(A) Classification of publications by major research area/category: Marker Paper and Cohort Descriptions (Marker and Cohort); Governance, Ethics, Community Engagement, and Biobanking (Ethics and Governance); Epidemiological studies on diseases, traits and risk factors (Epidemiology); Disease-associated behavior and awareness studies (Behavior and Awareness); Bioinformatics and Genomics Capacity Development (Bioinformatics); Genetics and Genomics Studies; Microbiome and Pathogen Studies. (B) Yearly breakdown of publications in each category. A few of these publications were aligned to more than one of the above-mentioned categories. Although we have cited these papers in each respective section, they have only been included in the category they best aligned to.
3 Marker Paper and Cohort Descriptions

The H3Africa marker paper, published in 2014 with 122 citations in PubMed Central articles, described the vision and anticipated trajectory of the Consortium emphasizing the need to develop capacity for health-related genomics research in Africa (Rotimi et al., 2014). In terms of the proposed measures of success, H3Africa has fully achieved success in 4 areas (high impact publications; establishing a pan-African bioinformatics network; establishing biorepositories; obtaining extended funding for another 5 years) and partially succeeded in the remaining 4 areas (increased availability of funding for African research; contributing to reversing the African brain-drain; regular and effective release of data; and the storage and release of samples). Several H3Africa Collaborative Centres, networks and projects have published their own marker papers or resource papers.

The H3Africa Consortium Pan-African Bioinformatics Network (H3ABio-Net) has published two landmark papers on sustainability and on the model for developing bioinformatics infrastructure and genomics research on the African continent (Mulder et al., 2016; 2017). The Collaborative African Genomics Network (CAfGEN) published a paper describing their network and explaining its objective of applying genomic technologies to probe host factors important in the progression of HIV and HIV-tuberculosis (TB) infection in sub-Saharan Africa (Mboowa et al., 2018).

Several of the projects involved in developing substantive African cohorts have published study-related marker/protocol papers and/or cohort descriptions, including the Stroke Investigative Research and Education Network (SIREN) (Akpalu et al., 2015; Owolabi et al., 2017; Adeoye, Oviagiele et al., 2017); the Africa Wits-INDEPTH Partnership for Genomic Studies (AWI-Gen) which is investigating the genetic and environmental contributions to cardiometabolic diseases in Africans (Ramsay et al., 2016; Ali et al., 2018); the African Collaborative Center for Microbiome and Genomics Research’s (ACC-ME’s) Human Papillomavirus (HPV) and Cervical Cancer Study (Adebamowo, Dareng et al., 2017); Kidney Disease Research Network (Osafo et al., 2016; 2015) and the H3Africa multi-center study on the prevalence and environmental and genetic determinants of type 2 diabetes in sub-Saharan Africa (Ekoru et al., 2016). These 5 studies, together with the rheumatic heart study, RHDGen, have harmonized and combined overlapping aspects of their data to develop the Cardiovascular H3Africa Innovation Resource (CHAIR) (Owolabi et al., 2019). CHAIR is expected to have >50,000 African participants from 13 different African countries, and a harmonized dataset with key cardiovascular disease (CVD)-related variables. In developing this collaborative resource, one of the main goals of the Consortium has been realized, which is to strengthen the
potential for novel discovery, to increase the power of analyses by combining resources, and to address research questions related to the health of Africans across the continent.

4 Governance, Ethics, Community Engagement and Biobanking

The success of any large and complex consortium rests on good governance and the development of sound and feasible policies and guidelines that support its research endeavour. One of the first challenges for the H3Africa Consortium was to develop policies and processes for genomic research and large data projects, with the long-term objective of sharing data and biospecimens with the global scientific community, for the purpose of improving the health of African and other global communities. This involves storing data in international repositories, in the case of H3Africa it is the European Genome-Phenome Archive, and developing biorepositories on the African continent. A series of workshops were held with ethics review committee members from many African institutions to examine the acceptance of a broad consent model for African Genomic Research; the equitable and fair sharing of data and samples; and the notion of benefit sharing (Ramsay et al., 2014; de Vries et al., 2016; Munung, Mayosi, and de Vries 2017; Tindana et al., 2019). Governance of digital health data in low and middle-income countries, biobank governance and potential exploitation of African biospecimens were considered in several publications (Staunton and Moodley 2013; Staunton and Moodley 2016; Tiffin, George, and Lefevre 2019). To better understand the African research landscape, there was an analysis on ethical guidelines, policies and procedures for research in 22 African countries (de Vries et al., 2017) and a governance framework was proposed for genomic research and biobanking in Africa (Yakubu et al., 2018).

Since African countries have different cultural norms and languages and their populations are mostly naïve to genomic research, much has been written by H3Africa research groups on appropriate African models or approaches for community engagement in the context of genomic research (Jenkins, Arulogun et al., 2016; Tindana et al., 2017; Pratt and de Vries 2018; Moodley and Beyer 2019; Singh et al., 2017; Staunton et al., 2019) the appropriate use of analogies and nomenclature, religious perspectives (Dennis-Antwi et al., 2019), return of incidental findings (Wonkam and de Vries 2020), the diagnostic/therapeutic misconceptions (Masiye, Mayosi, and de Vries 2017) and ethical and consent issues related to immortalized cell lines (de Vries et al., 2014; Campbell et al., 2018).
Consent models have been the topic of dedicated workshops, much discussion and publications since the H3Africa Consortium started its work with the intention of maximizing its resources by sharing data and biospecimens with the international research community. To enable future research through the use and re-use of data and biospecimens, it has been important to obtain broad consent from participants, either through a single consent or tiered consent model (Wright, Adeyemo, and Tiffin 2014; de Vries et al., 2016; Munung et al., 2016; Campbell, Susser et al., 2017; Tiffin 2018; Bukini et al., 2019). To ensure good governance, access to H3Africa resources is provided to the scientific community through a process of application to the H3Africa Data and Biospecimen Access Committee, and several research ethics committees have further stipulated that ethics approval is required for any new studies performed on these resources.

These activities have led to the development of three H3Africa guideline documents that are available from the H3Africa website and can be used by researchers who work in African communities when developing their own information sheets and informed consent documents: (1) H3Africa Guidelines for Community Engagement (2) H3Africa Guidelines for Informed Consent (3) H3Africa Guidelines for the Return of Individual Genetic Research Findings (https://h3africa.org).

H3Africa currently supports three African Biobanks, one each in Nigeria, Uganda and South Africa. Publications have addressed consent and stakeholder engagement for biobanking in Africa (de Vries et al., 2016; Moodley and Singh 2016; Staunton et al., 2018). Since building biobanks in resource limited environments is challenging, several papers have provided insights on building cost-effective biobanks and laboratory information management systems that align with international good practice (Abayomi et al., 2013; Matimba et al., 2016; Soo et al., 2017; Bendou et al., 2017; Ilboudo et al., 2017; Akinyemi et al., 2018; Abimiku et al., 2019). These publications have many practical suggestions on achieving ethically sound biobanking practices in resource-limited environments, with appropriate informed consent for the future use of samples.

5 Epidemiological Studies on Diseases, Traits and Risk Factors

An important outcome of the H3Africa Consortium is the development of large patient-control cohorts and population cross-sectional cohorts that provide an opportunity to assess the prevalence and distribution of cardiovascular and metabolic diseases and associated traits across the continent. In this section we describe 25 publications that examined and compared the distributions
of several traits and associated risk factors, including body mass index (BMI), obesity, stroke, carotid intima-media thickness (cIMT) and hypertension.

A series of publications described the BMI distribution and obesity in older adults in a population cross-sectional study by the AWI-Gen group, and included data from over 10,000 individuals sampled from six centres, both rural and urban, across four African countries, (Ramsay et al., 2016; Ramsay et al., 2018; Asiki et al., 2018; Haregu et al., 2018; Mikesfield et al., 2018; Wagner et al., 2018; Boua et al., 2018; Nonterah et al., 2018; Mohamed et al., 2019; Mashinya et al., 2018). These studies emphasised differences in the BMI distribution among East, West and South African populations, with higher BMI in the South and East, compared to the West (Ramsay et al., 2018). Differences in obesity were observed between men and women in the South and East, with women disproportionately affected, and over 65% of women from Soweto, South Africa, having a BMI >30. Dual-energy X-ray absorptiometry (DXA)-derived measures of fat distribution in Soweto further showed that women not only had higher BMI but also higher waist circumference (WC), fat mass (FM), subcutaneous adipose tissue (SAT), visceral adipose tissue (VAT), FM/fat free soft tissue mass (FFSTM) ratio and body fat percentage, in comparison to men (Pisa et al., 2018). Statistical models simultaneously compared the impact of various risk factors and demonstrated regional differences across Africa and between men and women. An in-depth characterization of several known BMI-associated risk factors, such as socio-economic status, diet and tobacco and alcohol use, highlighted potential explanations for the observed differences. For instance, while higher estimated socio-economic status (SES) was strongly associated with higher BMI across several of the African study sites, as reported for other low- and middle-income countries, this association was much stronger among women (Ramsay et al., 2018; Haregu et al., 2018). The association of BMI to partnership status, and level of education was also found to differ between the sexes and according to geographic region. Factors that were detected to contribute to lower BMI included smoking, drinking and infectious diseases such as TB and HIV (Ramsay et al., 2018). A key insight from these studies was the identification of specific associations between BMI and ethnicity, even among geographically and genetically proximal ethnic groups such as the Zulu and Tswana from South Africa (Mikesfield et al., 2018) and Kassena and Nankana from Ghana (Nonterah et al., 2018). These differences point to potential partitioning of genetic risk based on the characterization of fine-scale differences in genetic background.

Other CVD-associated traits that were investigated in the pan-African AWI-Gen cohort included cIMT, hypertension and chronic kidney disease. The distribution of cIMT, an early marker of atherosclerosis measured using
ultrasound, also showed strong regional variation, with the highest measurements in West African individuals, despite lower BMI levels (Nonterah et al., 2019). cIMT measurements were also found to demonstrate considerable sex differences, however, this was not uniform across the study sites. For example, women from Kenya were observed to have higher cIMT measurements compared to men whereas men in Burkina Faso demonstrated statistically significantly higher cIMT measurements compared to women. The major risk factors for higher cIMT identified in this study included increased age, BMI, systolic blood pressure, low-density lipoprotein cholesterol and smoking. High-density lipoprotein cholesterol (HDL-C), alcohol consumption and HIV were associated with lower cIMT.

The distribution of hypertension (defined as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg or taking anti-hypertensive medication) among older adult participants from the six study sites of the AWI-Gen study also showed striking regional differences in prevalence, awareness and control (Gómez-Olivé et al., 2017). The highest prevalence of hypertension was observed in Soweto (>50%) with several sites having significant differences between men and women. Sex-based comparisons detected a significantly higher prevalence of hypertension in women from Agincourt (South Africa), Dikgale (South Africa) and Nairobi (Kenya), while a reverse trend was observed for Nanoro (Burkina Faso). In a study on hypertension in a cohort of over 1000 adolescent sickle cell anaemia (SCA) patients from Tanzania, 8% had hypertension and risk factors that were independently associated with hypertension in SCA included age, BMI, pulse pressure and haemoglobin levels (Makubi et al., 2017).

Regional and sex differences were also observed for indices of kidney damage, including Estimated Glomerular Filtration Rate (eGFR) and albumin creatinine ratio (ACR), a measure of albuminuria. Chronic kidney disease (CKD), defined as eGFR<60 and ACR>3, had the highest prevalence in the South African study sites of AWI-Gen, being two-fold higher than in the West African study sites (George et al., 2019). Women were observed to have significantly higher prevalence compared to men. Major risk factors identified for kidney damage in this study included, age, hypertension, diabetes and HIV. Interestingly, about a third of the participants with CKD did not have diabetes, HIV infection or hypertension, suggesting the possible involvement of unidentified risk factors in African populations. An independent explorative study by the H3Africa Kidney Disease Research Network on the impact of perfluoralkyl substances (PFAS) (a class of compound used in household products that are associated with kidney dysfunction), demonstrated overall lower exposure to these compounds in African populations compared to American populations.
This study was based on 89 children and adolescents and showed that, in contrast to the American population-based National Health and Nutrition Examination Survey (NHANES) cohort that did not show any sex bias in the impact of PFASs, males in the African cohort were more strongly affected by the exposure compared to females.

A study aimed at characterising dyslipidaemia (hypercholesterolaemia, hypertriglyceridemia, elevated Low-density lipoprotein cholesterol (LDL-C) and low HDL-C) in 1839 individuals from the rural Ghana AWI-Gen study site (Novrongo) detected low HDL-C levels in about 60% of individuals (Agongo et al., 2018). Risk factors such as BMI, waist circumference and subcutaneous abdominal fat were found to be associated with different measures of dyslipidaemia in a sex-dependent manner. Notably, two of the well-known risk factors, age and diet, did not show any association with any form of dyslipidaemia, and SES was only found to be associated with HDL-C. As observed for BMI, the role of fine-scale ethnic differences (between Kassena and Nankana) was also observed for dyslipidaemia.

Phenotype data from the SIREN study, including 4200 individuals sampled across 15 study sites in Nigeria and Ghana, provided valuable estimates of the distribution of stroke and related risk factors in Central-West Africa (Owolabi et al., 2018). The primary risk factors associated with stroke included CVDs (hypertension, dyslipidaemia, diabetes and cardiac disease), diet (regular meat consumption, low green leafy vegetable consumption and added salt at the table), anthropometric indicators (elevated waist-to-hip ratio) and behavioural factors (stress, physical inactivity and current cigarette smoking) (Owolabi et al., 2018). Although some of these risk factors were also observed in a previous study (Feigin et al., 2016), the ranking of the contributions was found to be different in the SIREN cohort.

In attempts to identify biological parameters that could estimate the risk of stroke, the SIREN study investigated the role of cIMT as a risk factor using data from 555 participants from Nigeria. The results not only showed a strong association of cIMT with stroke but also demonstrated that it outperformed multivariable risk prediction schemes such as the Framingham’s Risk Score (FRS) and the Omnibus Risk Score (ORS) in estimating the overall risk of stroke (Owolabi, Akpa, and Agunloye 2016). Electrocardiographic profiles among 1020 acute stroke patients identified 90% with abnormal left ventricular (LV) geometry and 30% with systolic dysfunction, and showed that severe LV systolic dysfunction was significantly associated with one-month mortality (Adeoye et al., 2019). The high predictive value and regional variations in cIMT (Nonterah et al., 2019) call for more large-scale pan-African studies to introduce risk screening and timeous interventions.
A more nuanced analysis based on the classification of stroke by type, demonstrated a clear age stratification (Sarfo, Ovbiagele et al., 2018). In patients 50 years or younger, haemorrhagic and ischemic stroke were observed to be almost equally common, while ischemic stroke was three-fold higher in patients over 50 years (Sarfo, Ovbiagele et al., 2018). Further, classification of the type of stroke by sex detected that haemorrhagic stroke was more common among men, but with increased severity in women (Akpalu et al., 2019). Most risk factors, with the exception of salt intake and income, were shared by both men and women. Finally, ECG abnormalities also varied between stroke type and sex with atrial fibrillation being more common in women and ischemic stroke patients (Adeoye, Ogah et al., 2017).

Infectious diseases and specific monogenic traits related to malaria and HIV, both of which are common in Central-West Africa, have the potential to affect stroke outcome. A preliminary investigation of the impact of SCA and HIV on stroke outcome in a set of 35 patients from Nigeria demonstrated a worse 30-day mortality in patients with sickle cell trait (HbAS) compared to patients with normal adult haemoglobin (HbAA) (Olowoyo et al., 2016). In contrast, the impact of HIV, based on a study of 540 stroke cases and 540 controls from Ghana, showed that HIV infection was not associated with stroke outcome (Sarfo, Opare-Sem et al., 2018). However, due to the low incidence of HIV in West Africa (~2% in the study site), as well as the limited number of sickle cell patients in the cohort (n=35), these results need to be validated in larger cohorts. A study in Mali examined neurological complications in patients with sickle cell disease or trait, and showed that six out of eight patients with neurological symptoms had a diagnosis of stroke (Landouré et al., 2017). These studies highlight the need for further in-depth investigation of multi-morbidity in African populations.

6 Disease-Associated Behaviour and Awareness Studies

Planning and implementing approaches to address diseases at a population-scale requires careful assessment of the overall awareness of the diseases and the factors that increase risk. Although a plethora of studies have reported on these aspects in non-African populations, the varied socio-cultural settings in Africa have the potential to result in unique behavioural patterns that require careful investigation across countries and communities. In parallel to collecting phenotype information related to diseases, some of the H3Africa groups, based on questionnaires and interviews, have also collected valuable data on awareness, psychological and behavioural attitudes of individuals,
families and communities towards various diseases. This is well described in the reports from the SIREN group that provide key insights into awareness levels and attitudes towards stroke and stroke-associated risk factors (Ojagbemi et al., 2017; Sarfo, Nichols et al., 2017; Jenkins et al., 2018; Akinyemi et al., 2019). For example, independent surveys and interviews of stroke survivors from communities across Nigeria and Ghana show anxiety and stigma to be two major issues encountered by these patients (Jenkins, Arulogun, and Sarfo 2017; Ojagbemi et al., 2017; Akinyemi et al., 2019). The experience of stigma was found to be common among the stroke survivors interviewed, with 80% in a survey of 200 stroke survivors in Ghana reporting experiencing some form of stigma (Sarfo, Nichols et al., 2017). Interestingly, this was found to be more common among urban dwellers compared to rural, suggesting an important role of socio-cultural settings in this process. Similar investigations demonstrated about 20% of stroke survivors interviewed suffered from clinical anxiety and 70% of them were also found to suffer from depression (Ojagbemi et al., 2017). To investigate the source of the observed behavioural patterns, patient families and communities were interviewed, which showed clear gaps in awareness of the causes of stroke (Jenkins et al., 2018). Similarly, there was a major lack of understanding of the role of genetic susceptibility in stroke even among stroke survivors (57%) (Akinyemi et al., 2019). The studies strongly emphasize the need for developing comprehensive counselling approaches for patients and improving awareness at the community level for addressing these challenges.

A survey from the SIREN study also investigated the medication profile and adherence patterns in a set of stroke survivors in Ghana (Sarfo, Ovbiagele et al., 2017). The study showed the majority of patients to be on antihypertensive (94.5%), lipid-modifying (72.5%) and anti-platelet (65.6%) medication for the first twelve months following a stroke, and despite limited resources, 92% in this cohort remained on secondary prevention medications beyond a year. Other studies focused on the development of resources to enable efficient screening and intervention strategies, such as screening tools for stroke-associated depression (Ojagbemi et al., 2017b; 2017a), a monitoring tool for medication adherence (Jenkins, Burkett et al., 2016) and specialized training for healthcare professionals (Akinyemi et al., 2015). A major factor in the success of cohort surveys is the efficiencies with which the questionnaires capture data in a cohort. Three reports from the SIREN study describe the performance of different types of questionnaires in screening stroke/stroke free status among participants (Sarfo et al., 2016; Sarfo et al., 2016; 2017).

Publications involving the study of behavioural patterns included three studies from the ACCME group reporting on nuances and considerations in
Dareng et al. (2015) investigated the cultural barriers preventing women from going for cancer screenings and found that there were several misconceptions about the cause of cervical cancer as well as various cultural, religious and social factors that prevented women from going for cervical cancer screening (Dareng et al., 2015). Given the importance of particular patterns of sexual behaviour in the transmission of diseases, the ACCME group also examined the reliability of self-reporting of sexual behaviour history in a group of 725 urbanized women from Nigeria (Dareng et al., 2017). The study demonstrated overall high reliability in self-reporting of sexual behaviour and also emphasized the importance of interviewing skills and questionnaires, in capturing self-reported sexual history data reliably. Visual inspection with acetic acid (VIA) by health workers has been suggested as a low-cost solution for cervical cancer screening. However, a high element of subjectivity has been a major concern about this approach. The third study from the ACCME group, compared the outcomes from the VIA-based diagnosis to that from specialists, and highlighted the lack of interobserver concordance and objectivity in diagnosing cervical cancer using nurse-based VIA (Dareng et al., 2018).

Publications addressing attitudes and awareness towards diseases also included studies on hearing loss (Gardiner et al., 2019) and sickle-cell disease (Wonkam and Hurst 2014). An interview-based investigation of the content of delusion in 200 Xhosa-speaking individuals with schizophrenia, demonstrated a strong impact of cultural influence with a majority of participants believing bewitchment as the cause of their mental illness (Campbell, Sibeko et al., 2017). A questionnaire-based evaluation of the awareness of hypertension in the AWI-Gen cohort, found a generally low awareness of high blood pressure in rural and urban communities (Gómez-Olivé et al., 2017). Moreover, even among those who had been diagnosed, control was observed to be poor. The study also reported a considerable amount of sex-specific and regional variation in the awareness of hypertension. Although, some of these studies are relatively small in scale, as a whole they provide critical base-line data for further investigations and demonstrate the significance of cohort surveys in informing interventions, both at the level of the clinics and in informing public health policies.

7 Genetics and Genomics Studies

The genetics and genomics studies reported by the H3Africa Consortium can be broadly categorized into studies of monogenic diseases, and complex diseases and traits. The latter included candidate gene studies and genome-wide
association studies. In addition, there were population genetic studies. The monogenic disease studies predominantly addressed hearing impairment, neurological disorders and sickle cell anaemia.

Several studies, based on samples from Cameroon, South Africa and Ghana, examined the distribution of genetic variants for various forms of hearing impairment/deafness (HI) (Bosch et al., 2014; Wonkam et al., 2019; Adadey et al., 2019). These studies showed that while GJB2 gene variants are largely responsible for HI in European populations, the contribution of variants in this gene to HI in African populations is limited, highlighting the necessity of next-generation sequencing and functional genomics-based studies to identify genes and mutations responsible for HI in Africans (Bosch et al., 2014; Wonkam et al., 2019; Adadey et al., 2019). In addition, two recent publications reported on a large family with fragile-X syndrome (FXS) (Kamga et al., 2020) and MECP2 duplication syndrome (Tekendo-Ngongang et al., 2020) in families from Cameroon. The researchers also explored attitudes and understanding of genetic diseases in these settings. Another set of studies on neurological disorders reported cases of variation in the FA2H (Landouré et al., 2019), KIF5A (Guinto et al., 2017) and SPG11 (Landouré et al., 2020) genes causing hereditary spastic paraplegias (HSPs) in three independent families from Mali. Another study from this group reported a case study of a GARS mutation causing autosomal dominant Charcot-Marie-Tooth (CMT) syndrome in a consanguineous family from Mali (Yalcouyé et al., 2019).

Despite being a monogenic trait, the severity of sickle cell disease (SCD) and drug responses have been found to be modulated by a number of factors. This motivated studies to identify genetic variants underpinning various SCD related phenotypes (Gabriel and Przybylski 2010). Three candidate gene studies reported investigations into the genetic associations of SCD-related phenotypes in a cohort from Cameroon. The phenotypes analyzed in these studies included fetal hemoglobin levels (Pule et al., 2015; Pule, Bitoungui et al., 2017; Pule, Mnika et al., 2017), renal dysfunction (Geard et al., 2017), vaso-occlusive crises and the frequency of hospitalization (Wonkam et al., 2018). Genetic investigations into SCD also included a study on the impact of the drug hydroxyurea (HU) on miRNA expression in peripheral blood isolates of SCD patients (Mnika et al., 2019). Among the miRNAs detected to be differentially expressed following HU treatment, more than half were found to be associated with HbF regulatory genes (BCL11A, MYB, KLF-3, and SP1), indicating the possible significance of these genes in deciphering the mechanism of action of the drug and its therapeutic targets (Mnika et al., 2019).

The candidate gene studies for disease susceptibility included 7 publications related to trypanosomiasis, based on cohorts from Malawi, Cameroon,
Uganda, Guinea and Ivory Coast (Cooper et al., 2017; Ofon et al., 2017; Ahouty et al., 2017; Kaboré et al., 2017; Ofon et al., 2018; Kimuda et al., 2018; Kamoto et al., 2019). In addition to the apolipoprotein L1 (APOL1) trypanosomiasis protective variants, these studies also investigated the association of other genes with trypanosomiasis including IL10, IL8, IL4, HLAG, TNFA, TNX4LB, IL6, IFNG, MIF, APOL1, HLAA, IL1B, IL4R, IL12R, IL12RB, HP, HPR, and CFH. However, the studies were not uniform in design, with one focusing only on APOL1 variants (Cooper et al., 2017), while others included variants from 7 (Ofon et al., 2018), 8 (Kaboré et al., 2017), 16 (Ahouty et al., 2017), 17 (Ofon et al., 2017; Kamoto et al., 2019) and 18 genes (Kimuda et al., 2018). A key insight from these studies was the characterization of the complex and geographically dependent interaction of the APOL1 G1 and G2 variants with the two major Trypanosoma subtypes. For example, while neither of the APOL1 variants were protective against infection with the West African parasite, Trypanosoma brucei gambiense (Tbg), the APOL1 G2 variant was found to provide more than five-fold greater protection against the East African parasite, Trypanosoma brucei rhodesiense (Tbr) (Cooper et al., 2017). Moreover, while the G1 variant was associated with a reduced number of parasites in the blood, thereby decreasing severity, the G2 variant increased the severity of the infection by Tbg (Cooper et al., 2017). Therefore, from an evolutionary perspective, populations from East Africa that are exposed to both Tbg and Tbr, would have needed to develop a balance between the protective and potentially negative impacts of the G2 allele. Perhaps due to these complexities and despite its highly protective role and association to trypanosomiasis in populations from Uganda (Cooper et al., 2017) and Malawi (Kamoto et al., 2019), Kimuda et al. failed to detect any association of G2 with Tbr in an independent Ugandan population (Kimuda et al., 2018). The candidate gene studies also identified trypanosomiasis-associated variants in IL1A, (Ofon et al., 2018), IL1RN (Ofon et al., 2018) and IL6 (Kaboré et al., 2017) in different African populations. In addition, suggestive associations were detected for variants in IL6 and TNFA (Ahouty et al., 2017). Interestingly, two independent candidate gene studies for ischemic stroke in West Africans also converged on variants in APOL1 (Akinyemi et al., 2018) and IL6 (Voight et al., 2012), underlining the significance of these genes in both infectious and cardiovascular diseases.

Based on data generated on the Metabochip genotyping array, a specialized tool for replicating and fine-mapping variants associated with diseases and traits (including type 2 diabetes, coronary artery disease and myocardial infarction, body mass index, glucose and insulin levels, blood lipid levels, and blood pressure) (Voight et al., 2012), Sahibdeen et al. investigated the genetics of body composition in a South African cohort of 2000 participants (Sahibdeen et al., 2018). In addition to common anthropometric measurements such as height,
weight, hip and waist circumference, the study also tested the associations of quantitative traits based on DXA-derived measurements such as body fat percentage, fat mass, lean mass, visceral fat and subcutaneous fat. The study replicated signals around well-known loci such as $\text{FTO}$ (for waist to hip ratio), $\text{SEC16B}$ (for fat mass) and $\text{WARS2}$ (for waist to hip ratio) in the South African cohort (Sahibdeen et al., 2018). Another study conducted on the same cohort identified associations of variants in/around the $\text{NOS1AP}$, $\text{MYRF}$, $\text{POC1B}$, $\text{DACH1}$ and $\text{LPL}$ genes with blood pressure in South Africans (Hendry et al., 2018).

During the initial planning and design of projects under the H3Africa Consortium, the necessity for a cost-effective, Africa-centric genotyping array to enable efficient discovery of complex-trait associations, was recognized. To address this, the Consortium, in association with Illumina Inc, designed a custom ~2.3 million single nucleotide variant genotyping-array (Mulder et al., 2018). This array was optimized to capture common African variants, based on more than 3000 African whole genome sequences, and was enriched with genetic variants previously reported to be associated with various phenotypes. Moreover, to enable efficient genotype imputation for this array, the Consortium developed a reference panel and core imputation facility (Mulder et al., 2018). As the H3Africa genotyping array only became available in early 2018, most of the GWAS studies were delayed, explaining the scarcity of full-scale genome-wide associations studies in the current H3Africa publication list.

A survey of abstracts presented by groups at H3Africa Consortium and other international meetings/conferences suggests that genome-wide data analysis for many H3Africa studies are approaching completion, and two were recently published. The first was published in early 2020, reporting on an exome sequencing-based case-control study of schizophrenia in 1800 South Africans, which identified novel damaging variants in key synaptic function associated genes to be enriched in the cases (Gulsuner et al., 2020). The second was a study on the interaction of genetic variants and smoking on cIMT in 1776 West-African male participants, and also was the first H3Africa study based on genotype data generated on the H3Africa array (Boua et al., 2020). These developments predict major insights into the genetics of complex traits in African populations, from GWAS conducted by individual H3Africa groups as well as cross-Consortium studies such as CHAIR.

The characterization of intrinsic population structure and major demographic features, such as admixture and relatedness, are key to the effective design of genetic association studies in African populations. The high genetic diversity and unique linkage-disequilibrium (LD) architecture along with the scarcity of genome-scale data from some parts of the continent raised particular challenges for studying African populations. Although many of the
candidate gene studies have demonstrated extreme differences in the distribution of disease-associated alleles, based on only a limited number of variants, these studies did not provide suitable data for population-genetic research. The first study to investigate population structure at a genome-wide scale was a whole exome sequencing study of 314 children from Botswana and Uganda (Retshabile et al., 2018). In addition to demonstrating genomic distinctiveness of the Batswana population in comparison to 1000 Genomes Project populations, the study also identified a higher level of relatedness among the participants (Retshabile et al., 2018).

8 Bioinformatics and Genomics Capacity Development

Prior to the formation of the H3Africa Consortium, limited large scale genomic studies were being conducted on the African continent. As a result, there were few African-based researchers with the skills required to efficiently organize, store and analyse the large-scale biological datasets anticipated from the various H3Africa projects. With this in mind, the H3Africa Bioinformatics Network (H3ABioNet) was established to support H3Africa researchers through the development of bioinformatics capacity. H3ABioNet was established with over 30 nodes across 15 African countries with bioinformatics capacity ranging from established to intermediate, to little or no bioinformatics capacity. The initial overarching aim was to work towards building a network of nodes across Africa that would have the necessary skilled personnel and computational infrastructure to analyse the large H3Africa datasets. A major focus area of the initial activity for the network was core capacity development through the provision of teaching and training events, coupled with building core computational infrastructure at developing nodes (Table 9.1). This involved addressing some low-middle-income country (LMIC) specific challenges, such as limited/unstable network connectivity. Another major challenge was to monitor, follow up and advance the training imparted both at the participant and the node level, which was done through node accreditation in specific skills areas (e.g. GWAS and NGS workflows), internships and webinars. In addition, tools and resources were developed to facilitate communication between H3Africa projects and to ensure that a framework and policy guidelines were developed and implemented for the efficient organization, storage and analysis of the various datasets being generated (Table 9.1).

Building on this foundation, the emphasis of the network gradually shifted from capacity building to developing sustainable applied informatics solutions to continue to support the emerging requirements of the H3Africa Consortium.
| Challenge                      | Approach                                                                                                                                                                                                                                                                                                                                 |
|-------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Core infrastructure development | - Computational resources: 15 servers equipped with a total of 512 cores, 2384 GB RAM and 120 TB were installed  
                                - Training: hands on practical experience was provided with setting up and managing computing resources (Mulder et al. 2016)                                                                                                                                                              |
| Unstable internet connectivity | - Regular monitoring and troubleshooting of problematic network connections at each of the nodes  
                                - Globus Online service was used to facilitate the efficient transfer of large datasets over the internet  
                                - At sites identified to have limited internet connectivity, eBioKits (stand-alone devices containing pre-installed tools and databases) were set up (Mulder et al. 2016; Hernández-de-Diego et al. 2017) |
| Core capacity development     | - Extensive training: Over 30 face-to-face workshops were held from 2012 to 2019 on data management, system administration, genome-wide association studies (GWAS), next generation sequencing (NGS) data analysis, metagenomics, grant writing and professional development (Aron et al. 2017) |
| Follow up on capacity development | - An internship program was initiated to provide individuals with the opportunity to spend an extended period of time at a partner institute in order to develop specialized skills (Aron et al. 2017)  
                                - A webinar series was initiated to further encourage communication and collaboration between nodes (Fadlelmola et al. 2019)                                                                                          |
| Sustainable capacity development | - Bioinformatics course: a 3-month Introduction to Bioinformatics Training (IBT) online course, which has successfully run annually since 2016 (In 2016 alone, there were 364 participants across 20 classrooms in 10 African countries) (Gurwitz et al. 2017)  
                                - Curriculum development: The African Genomic Medicine Training (AGMT) course was developed that initially focused on a needs-assessment-based comprehensive genomics medicine curriculum, aimed, in the first instance, at nurses. The first course in 2017 included 225 participants in 19 classrooms across 11 countries, and it has been held annually since then (Nembaware and Mulder 2019).  
                                - Formation of an African Bioinformatics Education Committee tasked with developing a curriculum for setting up a bioinformatics degree program at African institutions (Mulder et al. 2016; Shaffer et al. 2019) |
This was achieved through advanced capacity development across the established network, setting up a data coordinating centre, providing high-quality informatics support, enabling and enhancing innovative translational research and continuing to build and promote the network beyond the Consortium to foster and promote external collaboration. To ensure that training activities are aligned with international standards and best practices in bioinformatics,
H3ABioNet worked closely with international training organisations such as the Global Organisation for Bioinformatics Learning, Education and Training (GOBLET), the European life-sciences Infrastructure for biological Information (ELIXIR) and the International Society for Computational Biology (ISCB). An Education Summit was held in May 2019 to bring together bioinformatics educators and trainers to refine core competencies and define guidelines on their application, and to develop additional bioinformatics training resources for the community. H3ABioNet has adopted and applied these competencies in the development of core training courses (Mulder et al., 2016; Mulder et al., 2018). Although still a challenging area, H3ABioNet has made great strides in building large-scale data analysis capacity on the continent.

Against the backdrop of a critical lack of bioinformatics expertise in Africa, a major accomplishment for the network was the development of a large number of resources and tools to support the analysis and exploration of data being generated by the larger H3Africa Consortium (Table 9.1). The tools include a novel post-GWAS approach called ancGWAS that was developed to improve the detection of disease variants with small effects by integrating the signals from a GWAS dataset, the local ancestry information and the SNP pairwise linkage disequilibrium in admixed populations (4-way complex admixture) into a protein-protein interaction network (Chimusa et al., 2016). Further exploration of the limitations of the current methods available in this area of post-GWAS analysis in admixed populations has been conducted (Chimusa et al., 2018; Geza et al., 2019; Awany, Allali, and Chimusa 2019; Chimusa et al., 2019) to identify the caveats and develop better suited methods to identify variants with small to moderate effects. The inability to access and use high-performance computing (HPC) clusters via a command line terminal is usually a limiting factor for most researchers wanting to analyse large datasets. Job Management System (JMS) was developed as a web-based front-end to an HPC cluster that allows users to create workflows based on different tools and to run, manage and monitor jobs on an HPC via an intuitive web-based graphical user interface (Brown et al., 2015). Several projects within the Consortium are centred around the identification and functional impact of African-specific genetic variants. The network also contributed to the development of specialized ontologies for sickle cell disease (Mulder et al., 2016; Adekile et al., 2019) and hearing impairment (Hotchkiss et al., 2019). A Human Mutation Analysis (HUMA) web server and database was designed to integrate sequence data, protein structure variation and disease data into a single connected database (Brown and Tastan Bishop 2018). HUMA allows for the uploading and interrogation of genetic variant data and the prediction of the impact at the protein structural level. A number of additional tools have been developed examining novel methods for modelling
the molecular dynamics of a protein including \textit{MD-TASK} and \textit{MODE-TASK} (Brown et al., 2017; Ross et al., 2018). On the pathogen front, Genome Detective is a web-based application that allows for the rapid assembly and identification of viral genomes from high throughput sequencing data and this tool has been used to accurately classify Dengue, Chikungunya and Zika viruses down to their species and sub-species levels (Vilsker et al., 2019; Fonseca et al., 2019).

9 Microbiome and Pathogen Studies

Several H3Africa projects are focusing on the contribution of the microbiome to disease susceptibility and progression in African settings. While a healthy microbiome promotes good health, changes in the microbiome can lead to severe disease, especially in younger children and immune-compromised individuals. In addition, outbreaks such as Ebola, Lassa Fever, Malaria and TB contribute significantly to the disease burden on the continent, especially in West Africa, and it is important to efficiently characterise circulating and novel viral pathogens using current and affordable technologies to reduce the impact of viral outbreaks (Folarin et al., 2016). Furthermore, understanding the host-pathogen interactions could lead to advancements in clinical care and treatment.

The Respiratory Microbiota of African Children (ReMac) Center has three projects aimed at describing the nasopharyngeal (NP) and upper airways microbiota and understanding the impact of lower respiratory tract infection and environmental exposures on the nasal microbiota in African children. \textit{Streptococcus pneumoniae} is the major bacterial cause of upper respiratory infections such as pneumonia in children under the age of 5 with a particularly high incidence in Africa (Rudan et al., 2013). Two independent longitudinal studies report on aspects of antimicrobial resistance to pneumococcal immunization with the pneumococcal conjugate vaccine (PCV13) aimed at reducing the colonization of the NP airways by 13 pneumococcal serotypes in children (Dube et al., 2018; Manenzhe et al., 2019). Similar investigations into carriage patterns of \textit{Staphylococcus aureus}, also known to colonise the NP epithelial surfaces, showed that it is a risk factor for a variety of other infections (Abdulgader et al., 2019). The nasal microbiome of children with pulmonary TB was also studied (Dube et al., 2016). The NP microbiome was shown to be influenced by environmental factors such as air pollution and tobacco smoke, and perturbation of the diversity of NP bacteria could lead to the development of lower respiratory tract infections. Similarly, another study aimed at examining the effect of indoor air pollution and environmental tobacco smoke on the NP
microbiome (in a cohort of mother-infant pairs during the first 12 months of life) showed exposure to antenatal environmental tobacco smoke to be associated with *Streptococcus pneumoniae* carriage in mothers, while postnatal tobacco smoke exposure was associated with carriage in infants (Vanker et al., 2019). Moreover, postnatal air pollution exposure was also associated with NP carriage of *Haemophilus influenzae* or *Moraxella catarrhalis* in infants. Therefore, exposure to environmental factors of both the mother and infant could result in an increased risk of lower respiratory tract infections.

The African Collaborative Center for Microbiome and Genomics Research (ACCME) aims to study the relationship between human papillomavirus (HPV), the vaginal microbiome and cervical cancer (Adebamowo, Dareng et al., 2017). Initial publications from this group assessed the incidence of cancer and cancer-associated infections in two Nigerian cancer registries, based on information collected between 2012 and 2014. A study examining the incidence of cancers attributed to infections revealed that 24% of cancers were associated to infections, while 22% were attributable to infections with the most common infectious agents being EBV, HPV, Hepatitis B and C, HIV and HHV8 (Odutola et al., 2016). A subsequent ACCME study confirmed that HPV infection was associated with a significant proportion of cancer cases in Nigerian women (Jedy-Agba et al., 2016). To further explore the prevalence of HPV, the age prevalence of HPV infection was determined in a sample of 278 women who presented for cervical cancer screening in Abuja, Nigeria. Based on questionnaire data, demographic characteristics, risk factors for cervical cancer and HPV genotyping using DNA extracted from cervical cells, the prevalence of HPV was detected to be 37% with the most prevalent type being HPV 35. HPV infection in women under 30 years was higher (52%) than those women over the age of 45 years (23%), showing a linear association between age and the prevalence of HPV infection (Akarolo-Anthony et al., 2014).

The persistence of high-risk HPV (hrHPV) infection plays a major role in the incidence and progression of cervical cancer and it is proposed that the vaginal microbiota play an important role in the persistence of hrHPV infection. Studies examining the vaginal microbiota and prevalent hrHPV infection in women in Nigeria found a suggestive association between prevalent hrHPV infection and a reduced abundance of *Lactobacillus sp.* and an increased abundance of the genera *Prevotella* and *Leptotrichia* in HIV-negative women (Dareng et al., 2016). Possible association between *Mycoplasma hominis* and persistent hrHPV in the vaginal microbiota was also reported (Adebamowo, Ma et al., 2017). In addition, significant associations between prevalent and persistent hrHPV infections and specific HLA haplotypes were detected, suggesting a possible genetic risk factor for hrHPV infection in African women.
(Adebamowo and Adeyemo 2019). Analysis of the role of HIV infection in cervical cancer in women indicates possible interactions between HIV and hrHPV (Adebamowo, Olawande et al., 2017; Adebamowo et al., 2018). Furthermore, association between cervical HPV11, HIV and genital warts was also observed in women in Nigeria (Dareng et al., 2019). These studies highlight the burden of HPV infection in cervical cancer and the role of potential screening and other intervention strategies such as vaccines to reduce the incidence of cervical cancer in African women.

Infectious diseases contribute significantly to the disease burden in Africa. In particular, West African countries are exposed to periodic outbreaks of viral diseases such as Ebola and Lassa Fever. A number of H3Africa studies have focused on sequencing-based in-depth analysis of these pathogens to generate insights into infection outbreaks. These include studies of Ebola pathogens (Folarin et al., 2016), RNA viruses (Stremlau et al., 2015) and Lassa virus genomes (Siddle et al., 2018). A study of RNA viruses in unexplained acute febrile disease (UAFI) patients and healthy individuals in a Nigerian community revealed the presence of many well characterized viruses in the blood of the UAFI patients (Stremlau et al., 2015). These included HIV-1, hepatitis B and C and Lassa virus. The study also identified two novel rhabdoviruses, Ekpoma virus 1 (EKV-1) and Ekpoma virus 2 (EKV-2), isolated from two healthy female individuals, which are similar to the Bas-Congo virus identified in a patient with viral haemorrhagic fever. Further analysis revealed exposure to EKV-1 and EKV-2 in a larger healthier cohort at 10% and 50% respectively (Stremlau et al., 2015).

While the incidence of malaria infection and the mortality rate has declined in most African countries, it is still a major health threat, especially in certain parts of the continent where malaria is endemic. The identification and understanding of the genomic diversity of Plasmodium falciparum have been explored using various approaches, however, a standardized method is still needed to monitor population dynamics, transmission and drug resistance between closely related parasites. A study utilizing the malaria barcode, which is a 24 single nucleotide polymorphism based molecular barcode assay, was able to differentiate between closely related P. falciparum infections in two urban cities in Nigeria. The results showed a low level of intra-population diversity in P. falciparum and a low degree of polygenomic infections across the two groups that had not been observed before. This indicates that the 24-SNP barcode method is efficient at monitoring changes in parasite population diversity and divergence over time and can be extended to explore transmission patterns and drug resistance in future studies (Bankole et al., 2018).
Tuberculosis is another major illness that impacts millions of people in Africa. A number of studies based on an Ethiopian cohort, generated insights into aspects of tuberculosis infection, diagnosis and treatment. Studies on *Mycobacterium tuberculosis* complex (MBTC) diversity (Nuru et al., 2015; Bedewi et al., 2017) showed significant geographic diversity of bacterial strains within a country and reported several strains that were not present in the molecular genotyping databases of MTBC. As the accurate diagnosis of TB and latent TB infection (LTBI) play an integral role in administering the correct treatment regime to tackle the infection, the detection of LTBI plays a key role in addressing the disease. Culture-based methods are routinely used to diagnose of active TB, but have limited sensitivity and efficiency. A study reporting an antibody microarray-based approach for studying abundance of cytokine and chemokine showed an increase in IFN-γ and interleukin 17 levels that might serve as good indicators of LTBI (Teklu et al., 2018). Moreover, the chemokines RANTES and MIP-1β showed the potential for use in differentiating between combined active TB and LTBI groups, and the unaffected control group (Teklu et al., 2018). Similarly, analysis of the performance of the GenoType MTBDRplus assay, an assay aimed at investigating the development of multidrug-resistant strains (MDR-TB) also demonstrated the possibility of improving both efficiency and reducing the time needed to diagnose MDR-TB (Bedewi et al., 2016). Two studies that aimed to estimate the distribution of MDR-TB in Ethiopia showed a relatively lower number of MDR isolates in new TB cases, compared to previous studies hinting at a positive impact of current intervention measures (Bedewi et al., 2017; Wondale et al., 2018; Alelign et al., 2019).

Epilogue - H3Africa Data in Population-Genetics Research

Although the H3Africa Consortium's main objective is to identify genetic underpinnings of communicable and non-communicable diseases in Africa, the data generated for these studies are playing an instrumental role in generating insights into the population structure and demographic history of African populations. Two key studies pertinent to African population genetics were published just after the period under review, and are briefly discussed because of their relevance to the theme of this book.

The first is a South African study based on data from the AWI-Gen project (Sengupta et al., 2021). It demonstrates the valuable role of consortium data in enhancing our understanding of the genetics of a particular country or geographic region. This population-cross sectional dataset, generated primarily...
for studying cardio-metabolic diseases (Ramsay et al., 2016), included most of the major South African Bantu-speaking groups, was used to investigate population structure. The analyses of genome-wide genetic data demonstrated a clear population structure within the South-Eastern Bantu-speaking (SEB) ethnolinguistic groups in the country. The structure showed correspondence with geographic location and languages and provided insights into the migration and admixture events that may have contributed to the current distribution of SEB groups across the country. Moreover, by recording parental and grandparental ethnicity it became apparent that recent inter-group admixture has made the study of population structure more complex and that this would need to be considered when performing disease-association studies.

The second study is based on whole-genome sequence data that was generated primarily for the design of a genotyping array for the Consortium, and resulted in a landmark population genetics paper for H3Africa. This pan-African study included participants from 50 ethnolinguistic groups sampled across 13 countries and provided the most comprehensive description of genetic diversity and population structure across African populations (Choudhury et al., 2020). Despite the recent inclusion of variants from thousands of African genomes in current databases, the study identified over 3 million novel variants from only 300 genomes. Some of the novel findings included the detection of gene flow events that identified East African gene flow into a population in Nigeria, and rain-forest forager gene flow into a population from Uganda. This highlights the potential of genetic studies to detect major differences in genetic and demographic histories of geographically neighbouring populations. These differences are also pertinent to the distribution of disease-causing variants such as the sickle cell mutation (HbS). While supporting the overlap in the distribution of HbS and malaria in Africa, this study noted marked differences in HbS allele frequency in two presently neighbouring but historically distant populations from Uganda. A study based on a subset of the data further explored the genetic diversity of the Nilo-Saharan populations and their genetic contributions to the Niger-Congo populations from neighbouring regions (Mulindwa et al., 2020).

These studies underline the promise of novel population genetic insights from genomic data generated by the H3Africa Consortium projects, and emphasise the need for including previously understudied geographic regions and ethnic groups. The aggregation and in-depth analysis of the genotype-array and whole-genome sequence data being generated by H3Africa studies promise to provide robust insights into genetic diversity and demographic history of African ethnolinguistic groups. Moreover, African-centric resources such as a genotyping array, imputation panel, local ancestry reference panels and...
variation databases are being developed by the Consortium and are expected to help researchers to capture and analyse African genetic data at a greater depth.

11 Conclusions

Over a period of seven years the publications from the H3Africa Consortium have shown a significant shift from reviews and perspectives to original research papers that contribute new knowledge to understanding health and disease in Africa (Figure 9.4B). Although presently skewed toward phenotype and behavioural studies, many genetics and genomics papers have been published and are expected to increase as several projects analyse and interpret their genomic data. Key words from the titles of the 169 core H3Africa papers have been assessed and quantified to highlight the most active themes under six domain categories (Figure 9.5). Once H3Africa Consortium project data are submitted to the European Genome-Phenome archive and access is provided by the H3Africa Data and Biospecimens Access Committee the number of publications that use these resources for discovery, validation and comparative purposes will increase. We anticipate that the next three years will be characterized by more complex epidemiological models to tease out disease mechanisms, and by genomic studies to identify genetic variants associated with disease risk. The genome-wide genotype data and whole-genome sequence data will be available to the research community to further explore population genetics studies to reveal hidden population demographic histories and to compare these findings to hypotheses put forward by linguistic and anthropological research. In addition, the Consortium will contribute to developing African-appropriate approaches to data analysis and to a better understanding of the value that African genetic diversity, together with rich phenotype, behavioural and infection data, can bring to a global understanding of health and disease.

Acknowledgements

H3Africa is primarily funded by the US National Institutes of Health and the UK Wellcome Trust, with support from the African Academy of Sciences. The authors are members of the AWI-Gen Collaborative Centre funded by the National Human Genome Research Institute (NHGRI), Office of the Director (OD), Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD), the National Institute of Environmental Health Sciences (NIEHS), the Office of AIDS Research (OAR) and the National
Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), of the National Institutes of Health (NIH) under award number U54HG006938 and its supplements, as part of the H3Africa Consortium. MR is a South African Research Chair in Genomics and Bioinformatics of African populations hosted by the University of the Witwatersrand, funded by the Department of Science and Technology, and administered by the National Research Foundation. AC
and DS were supported by the AWI-Gen grant. SA is a member of H3ABioNet funded by the US National Institutes of Health Common Fund Grant Numbers U41HG008941 and U24HG009641. We sincerely thank Michelle Skelton and her team at the H3Africa Coordinating Centre for the map in Figure 9.1 and the latest statistics on the consortium and Harry Wedel for sharing the search terms for the PubMed search.

References

Publications other than core H3Africa publications are indicated in blue

Abayomi, A., Christoffels, A., Grewal, R., Karam, L.A., Rossouw, C., Staunton, C., Swanepoel, C., and Rooyen, B. van. 2013. Challenges of biobanking in South Africa to facilitate indigenous research in an environment burdened with human immunodeficiency virus, tuberculosis, and emerging noncommunicable diseases. *Biopreservation and Biobanking* **11**(6): 347–354. https://doi.org/10.1089/bio.2013.0049.

Abdulgader, S.M., Robberts, L., Ramjitth, J., Nduru, P.M., Dube, F., Gardner-Lubbe, S., Zar, H.J., and Nicol, M.P. 2019. Longitudinal population dynamics of *Staphylococcus aureus* in the nasopharynx during the first year of life. *Frontiers in Genetics* **10**(3): 1–10. https://doi.org/10.3389/fgene.2019.00198.

Abimiku, A.G., Croxton, T., Ozumba, P.J., Agala, N., Balogun, O., Jonathan, E., Onyemata, E., et al. 2019. Blueprint for building a biorepository in a resource-limited setting that follows international best practices. *African Journal of Laboratory Medicine* **8**(1): 1–12. https://doi.org/10.4102/ajlm.v8i1.722.

Adadey, S.M., Manyisa, N., Mnika, K., Kock, C. de, Nembaware, V., Quaye, O., Amedofu, G.K., Awandare, G.A., and Wonkam, A. 2019. GJB2 and GJB6 mutations in non-syndromic childhood hearing impairment in Ghana. *Frontiers in Genetics* **10**(9): 1–8. https://doi.org/10.3389/fgene.2019.00841.

Adebamowo, S.N., and Adeyemo, A.A. 2019. Classical HLA alleles are associated with prevalent and persistent cervical high-risk HPV infection in African women. *Human Immunology* **80**(9): 723–730. https://doi.org/10.1016/j.humimm.2019.04.011.

Adebamowo, S.N., Dareng, E.O., Famooto, A.O., Offiong, R., Olaniyan, O., Obende, K., Adebayo, A., et al. 2017. Cohort Profile: African collaborative center for microbiome and genomics research’s (ACCME’s) human papillomavirus (HPV) and cervical cancer study. *International Journal of Epidemiology* **46**(6): 1745–1745j. https://doi.org/10.1093/ije/dyx050.

Adebamowo, S.N., Famooto, A., Dareng, E.O., Olawande, O., Olaniyan, O., Offiong, R., and Adebamowo, C.A. 2018. Clearance of type-specific, low-risk, and high-risk cervical human papillomavirus infections in HIV-negative and HIV-positive women. *Journal of Global Oncology* **4**(July): 1–12. https://doi.org/10.1200/JGO.17.00129.
Adebamowo, S.N., Ma, B., Zella, D., Famooto, A., Ravel, J., and Adebamowo, C. 2017. *Mycoplasma hominis* and *Mycoplasma genitalium* in the vaginal microbiota and persistent high-risk human papillomavirus infection. *Frontiers in Public Health* 5(June): 1–10. https://doi.org/10.3389/fpubh.2017.00140.

Adebamowo, S.N., Olawande, O., Famooto, A., Dareng, E.O., Offiong, R., and Adebamowo, C.A. 2017. Persistent low-risk and high-risk human papillomavirus infections of the uterine cervix in HIV-negative and HIV-positive women. *Frontiers in Public Health* 5(July): 1–11. https://doi.org/10.3389/fpubh.2017.00178.

Adekile, A., Anie, K.A., Hamda, C. ben, Brown, B., Bukini, D., Campbell, A., Chaouch, M., et al. 2019. The sickle cell disease ontology: Enabling universal sickle cell-based knowledge representation. *Database* 2019(November): 1–12. https://doi.org/10.1093/database/baz118.

Adeoye, A.M., Ogah, O.S., Ovbiagele, B., Akinyemi, R., Shidali, V., Agyekum, F., Aje, A., et al. 2017. Prevalence and prognostic features of ECG abnormalities in acute stroke: Findings from the SIREN study among Africans. *Global Heart* 12(2): 99–105. https://doi.org/10.1016/j.gheart.2017.01.002.

Adeoye, A.M., Ovbiagele, B., Akinyemi, J.O., Ogah, O.S., Akinyemi, R., Gebregziabher, M., Wahab, K., et al. 2019. Echocardiographic abnormalities and determinants of 1-month outcome of stroke among West Africans in the SIREN Study. *Journal of the American Heart Association* 8(11): 1–11. https://doi.org/10.1161/JAHA.118.010814.

Adeoye, A.M., Ovbiagele, B., Kolo, P., Appiah, L., Aje, A., Adebayo, O., Sarfo, F., et al. 2017. Exploring overlaps between the genomic and environmental determinants of LVH and stroke: A multicenter study in West Africa. *Global Heart* 12(2): 107–113.e5. https://doi.org/10.1016/j.gheart.2017.01.001.

Agongo, G., Nontereha, E.A., Debpuur, C., Amenga-Etego, L., Ali, S., Oduro, A., Crowther, N.J., and Ramsay, M. 2018. The burden of dyslipidaemia and factors associated with lipid levels among adults in rural northern Ghana: An AWI-Gen sub-study. *PLoS One* 13(11): 1–21. https://doi.org/10.1371/journal.pone.0206326.

Ahoutry, B., Koffi, M., Ilboudo, H., Simo, G., Matovu, E., Mulindwa, J., Hertz-Fowler, C., et al. 2017. Candidate genes-based investigation of susceptibility to human African trypanosomiasis in Côte d’Ivoire. *PLoS Neglected Tropical Diseases* 11(10): 1–13. https://doi.org/10.1371/journal.pntd.0005992.

Akarolo-Anthony, S.N., Famooto, A.O., Dareng, E.O., Olaniyan, O.B., Offiong, R., Wheeler, C.M., and Adebamowo, C.A. 2014. Age-specific prevalence of human papillomavirus infection among Nigerian women. *BMC Public Health* 14(1): 1–7. https://doi.org/10.1186/1471-2458-14-656.

Akinyemi, R., Tiwari, H.K., Arnett, D.K., Ovbiagele, B., Irvin, M.R., Wahab, K., Sarfo, F., et al. 2018. APOL1, CDKN2A/CDKN2B, and HDAC9 polymorphisms and small vessel ischemic stroke. *Acta Neurologica Scandinavica* 137(1): 133–141. https://doi.org/10.1111/ane.12847.
Akinyemi, R.O., Akinwande, K., Diala, S., Adeleye, O., Ajose, A., Issa, K., Owusu, D., et al. 2018. Biobanking in a challenging African environment: Unique experience from the SIREN Project. Biopreservation and Biobanking 16(3): 217–232. https://doi.org/10.1089/bio.2017.0113.

Akinyemi, R.O., Owolabi, M.O., Adebayo, P.B., Akinyemi, J.O., Otubogun, F.M., Uvere, E., Adeniji, O., et al. 2015. Task-shifting training improves stroke knowledge among Nigerian non-neurologist health workers. Journal of the Neurological Sciences 359(1–2): 112–116. https://doi.org/10.1016/j.jns.2015.10.019.

Akinyemi, R.O., Sarfo, F.S., Akinyemi, J., Singh, A., Onoja Akpa, M., Akpalu, A., Owolabi, L., et al. 2019. Knowledge, attitudes and practices of West Africans on genetic studies of stroke: Evidence from the SIREN Study. International Journal of Stroke 14(1): 69–79. https://doi.org/10.1177/1747493018790059.

Akpalu, A., Gebregziabher, M., Ovbiagele, B., Sarfo, F., Iheonye, H., Akinyemi, R., Akpa, O., et al. 2019. Differential impact of risk factors on stroke occurrence among men versus women in West Africa: The SIREN Study. Stroke 50(4): 820–827. https://doi.org/10.1161/STROKEAHA.118.022786.

Akpalu, A., Sarfo, F.S., Ovbiagele, B., Akinyemi, R., Gebregziabher, M., Obiako, R., Owolabi, L., et al. 2015. Phenotyping stroke in sub-Saharan Africa: Stroke investigative research and education network (SIREN) Phenomics Protocol. Neuroepidemiology 45(2): 73–82. https://doi.org/10.1159/000437372.

Alelign, A., Zewude, A., Mohammed, T., Tolosa, S., Ameni, G., and Petros, B. 2019. Molecular detection of Mycobacterium tuberculosis sensitivity to rifampicin and isoniazid in South Gondar Zone, northwest Ethiopia. BMC Infectious Diseases 19(1): 1–8. https://doi.org/10.1186/s12879-019-3978-3.

Ali, S.A., Soo, C., Agongo, G., Alberts, M., Amenga-Etego, L., Boua, R.P., Choudhury, A., et al. 2018. Genomic and environmental risk factors for cardiometabolic diseases in Africa: methods used for Phase 1 of the AWI-Gen population cross-sectional study. Global Health Action 11(sup2): 1–11. https://doi.org/10.1080/16549716.2018.1507133.

Aron, S., Gurwitz, K., Panji, S., Mulder, N., and Consortium, H. 2017. H3ABioNet: Developing Sustainable Bioinformatics Capacity in Africa. EMBnet Journal 23(3): 886. https://doi.org/10.14806/ej.23.3.886.

Asiki, G., Mohamed, S.F., Wambui, D., Wainana, C., Muthur, S., Ramsay, M., and Kyobutungi, C. 2018. Sociodemographic and behavioural factors associated with body mass index among men and women in Nairobi slums: AWI-Gen Project. Global Health Action 11(sup2): 1–11. https://doi.org/10.1080/16549716.2018.1470738.

Awany, D., Allali, I., and Chimusa, E.R. 2019. Tantalizing dilemma in risk prediction from disease scoring statistics. Briefings in Functional Genomics 18(4): 211–219. https://doi.org/10.1093/bfgp/ely040.
Baichoo, S., Soulimi, Y., Panji, S., Botha, G., Meintjes, A., Hazelhurst, S., Bendou, H., et al. 2018. Developing reproducible bioinformatics analysis workflows for heterogeneous computing environments to support African genomics. BMC Bioinformatics 19(1): 1–13. https://doi.org/10.1186/s12859-018-2446-1.

Bankole, B.E., Kayode, A.T., Nosamiefan, I.O., Eromon, P., Baniecki, M.L., Daniels, R.F., Hamilton, E.J., et al. 2018. Characterization of Plasmodium falciparum structure in Nigeria with malaria SNPs barcode. Malaria Journal 17(1): 1–10. https://doi.org/10.1186/s12936-018-2623-8.

Bedewi Omer, Z., Mekonnen, Y., Worku, A., Zewde, A., Medhin, G., Mohammed, T., Pieper, R., and Ameni, G. 2016. Evaluation of the GenoType MTBDRplus assay for detection of rifampicin- and isoniazid-resistant Mycobacterium tuberculosis isolates in central Ethiopia. International Journal of Mycobacteriology 5(4): 475–481. https://doi.org/10.1016/j.ijmyco.2016.06.005.

Bedewi, Z., Mekonnen, Y., Worku, A., Medhin, G., Zewde, A., Yimer, G., Pieper, R., and Ameni, G. 2017. Mycobacterium tuberculosis in central Ethiopia: drug sensitivity patterns and association with genotype. New Microbes and New Infections 17(May): 69–74. https://doi.org/10.1016/j.nmni.2017.02.003.

Bedewi, Zufan, Worku, A., Mekonnen, Y., Yimer, G., Medhin, G., Mamo, G., Pieper, R., and Ameni, G. 2017. Molecular typing of Mycobacterium tuberculosis complex isolated from pulmonary tuberculosis patients in central Ethiopia. BMC Infectious Diseases 17(1): 2–9. https://doi.org/10.1186/s12879-017-2267-2.

Bendou, H., Domelevo Entfellner, J.B., Heusden, P. van, Gamieldien, J., and Tiffin, N. 2016. NetCapDB: Measuring bioinformatics capacity development in Africa. BMC Research Notes 9(1): 1–8. https://doi.org/10.1186/s13104-016-1950-5.

Bendou, H., Sizani, L., Reid, T., Swanepoel, C., Ademuyiwa, T., Merino-Martinez, R., Meuller, H., Abayomi, A., and Christoffels, A. 2017. Baobab laboratory information management system: Development of an open-source laboratory information management system for Biobanking. Biopreservation and Biobanking 15(2): 116–120. https://doi.org/10.1089/bio.2017.0014.

Bosch, J., Noubiap, J.J.N., Dandara, C., Makubalo, N., Wright, G., Entfellner, J.-B.D., Tiffin, N., and Wonkam, A. 2014. Sequencing of GJB2 in Cameroonian and black South Africans and comparison to 1000 Genomes Project data support need to revise strategy for discovery of nonsyndromic deafness genes in Africans. Omics: A Journal of Integrative Biology 18(11): 705–710. https://doi.org/10.1089/omi.2014.0063.

Boua, P.R., Brandenburg, J.T., Choudhury, A., Hazelhurst, S., Sengupta, D., Agongo, G., Nonterah, E.A., et al. 2020. Novel and known gene-smoking interactions with cIMT identified as potential drivers for atherosclerosis risk in West-African populations of the AWI-Gen study. Frontiers in Genetics. 10(February): 1–21 https://doi.org/10.3389/fgene.2019.01354.
Choudhury et al. 2018. Gender differences in sociodemographic and behavioural factors associated with BMI in an adult population in rural Burkina Faso—an AWI-Gen sub-study. *Global Health Action* 11(sup2): 34–44. https://doi.org/10.1080/16549716.2018.1527557.

Brown, D.K., Penkler, D.L., Amamuddy, O.S., Ross, C., Atilgan, A.R., Atilgan, C., and Bishop, Ö.T. 2017. MD-TASK: A software suite for analyzing molecular dynamics trajectories. *Bioinformatics* 33(17): 2768–2771. https://doi.org/10.1093/bioinformatics/btx349.

Brown, D.K., Penkler, D.L., Musyoka, T.M., and Bishop, Ö.T. 2015. JMS: An open source workflow management system and web-based cluster front-end for high performance computing. *PLoS One* 10(8): 1–25. https://doi.org/10.1371/journal.pone.0134273.

Brown, D.K., and Tastan Bishop, Ö. 2018. HUMA: A platform for the analysis of genetic variation in humans. *Human Mutation* 39(1): 40–51. https://doi.org/10.1002/humu.23334.

Bukini, D., deVries, J., Treadwell, M., Anie, K., Dennis-Antwi, J., Kamga, K.K., McCurdy, S., Ohene-Frempong, K., Makani, J., and Wonkam, A. 2019. Exploring the role of shared decision making in the consent process for pediatric genomics research in Cameroon, Tanzania, and Ghana. *AJOB Empirical Bioethics* 10(3): 182–189. https://doi.org/10.1080/23294515.2019.1645759.

Campbell, M.M., Sibeko, G., Mall, S., Baldinger, A., Nagdee, M., Susser, E., and Stein, D.J. 2017. The content of delusions in a sample of South African Xhosa people with schizophrenia. *BMC Psychiatry* 17(1): 1–9. https://doi.org/10.1186/s12888-017-1196-3.

Campbell, M.M., Susser, E., Mall, S., Mqulwana, S.G., Mndini, M.M., Ntola, O.A., Nagdee, M., Zingela, Z., Wyk, S. van and Stein, D.J. 2017. Using iterative learning to improve understanding during the informed consent process in a South African psychiatric genomics study. *PLoS One* 12(11): 1–11. https://doi.org/10.1371/journal.pone.0188466.

Campbell, M.M., Vries, J. de, Mqulwana, S.G., Mndini, M.M., Ntola, O.A., Jonker, D., Malan, M., et al. 2018. Predictors of consent to cell line creation and immortalisation in a South African schizophrenia genomics study. *BMC Medical Ethics* 19(1): 1–7. https://doi.org/10.1186/s12910-018-0333-2.

Chimusa, E.R., Dalvie, S., Dandara, C., Wonkam, A., and Mazandu, G.K. 2019. Post genome-wide association analysis: Dissecting computational pathway/network-based approaches. *Briefings in Bioinformatics* 20(2): 690–700. https://doi.org/10.1093/bib/bby035.

Chimusa, E.R., Defo, J., Thami, P.K., Awany, D., Mulisa, D.D., Allali, I., Ghazal, H., Moussa, A., and Mazandu, G.K. 2018. Dating admixture events is unsolved problem in multi-way admixed populations. *Briefings in Bioinformatics*, 21(October): 144–155. https://doi.org/10.1093/bib/bby112.
Chimusa, E.R., Mbiyavanga, M., Mazandu, G.K., and Mulder, N.J. 2016. AncGWAS: A post genome-wide association study method for interaction, pathway and ancestry analysis in homogeneous and admixed populations. *Bioinformatics* 32(4): 549–556. https://doi.org/10.1093/bioinformatics/btv619.

Choudhury, A., Aron, S., Botiguë, L.R., Sengupta, D., Botha, G., Bensellak, T., Wells, G., Kumuthini, J., Shriner, D., Fakim, Y.J., Ghoorah, A.W., Dareng, E., Odia, T., Falola, O., Adebiyi, E., Hazelhurst, S., Mazandu, G., Nyangiri, O.A., et al. 2020. High-depth African genomes inform human migration and health. *Nature* 586(7831): 741–748. https://doi.org/10.1038/s41586-020-2859-7.

Cooper, A., Ilboudo, H., Alibu, V.P., Ravel, S., Enyaru, J., Weir, W., Noyes, H., et al. 2017. APOL1 renal risk variants have contrasting resistance and susceptibility associations with African trypanosomiasis. *ELife* 6(May): e25461. https://doi.org/10.7554/eLife.25461.

Dareng, E.O., Ma, B., Famooto, A.O., Akarolo-Anthony, S.N., Offiong, R.A., Olaniyan, O., Dakum, P.S., et al. 2016. Prevalent high-risk HPV infection and vaginal microbiota in Nigerian women. *Epidemiology and Infection* 144(1): 123–137. https://doi.org/10.1017/S0950268815000965.

Dareng, E.O., Adebamowo, S.N., Eseyin, O.R., Odutola, M.K., Pharoah, P.P., and Adebamowo, C.A. 2017. Test-retest reliability of self-reported sexual behavior history in urbanized nigerian women. *Frontiers in Public Health* 5(July): 1–9. https://doi.org/10.3389/fpubh.2017.00172.

Dareng, E.O., Adebamowo, S.N., Famooto, A., Olawande, O., Odutola, M.K., Olaniyan, Y., Offiong, R.A., Pharoah, P.P., and Adebamowo, C.A. 2019. Prevalence and incidence of genital warts and cervical human papillomavirus infections in nigerian women 11 medical and health sciences 117 public health and health services. *BMC Infectious Diseases* 19(1): 1–10. https://doi.org/10.1186/s12879-018-3582-y.

Dareng, E.O., Jedy-Agba, E., Bamisaye, P., Isa Modibbo, F., Oyeneyin, L.O., Adewole, A.S., Olaniyan, O.B., Dakum, P.S., Pharoah, P.D., and Adebamowo, C.A. 2015. Influence of spirituality and modesty on acceptance of self-sampling for cervical cancer screening. *PLoS One* 10(11): 1–12. https://doi.org/10.1371/journal.pone.0141679.

Dareng, E.O., Olaniyan, Y., Odutola, M.K., Adebamowo, S.N., Famooto, A., Offiong, R., Obende, K., et al. 2018. Secular trend in interobserver agreement of VIA diagnosis for cervical cancer screening in Nigeria. *PLoS One* 13(12): 1–14. https://doi.org/10.1371/journal.pone.0208531.

Dennis-Antwi, J.A., Ohene-Frempong, K., Anie, K.A., Dzikunu, H., Agyare, V.A., Okyere Boadu, R., Sarfo Antwi, J., et al. 2019. Relation between religious perspectives and views on sickle cell disease research and associated public health interventions in Ghana. *Journal of Genetic Counseling* 28(1): 102–118. https://doi.org/10.1007/s10897-018-0296-7.
Dube, F.S., Kaba, M., Robberts, F.J.L., Tow, L.A., Lubbe, S., Zar, H.J., and Nicol, M.P. 2016. Respiratory microbes present in the nasopharynx of children hospitalised with suspected pulmonary tuberculosis in Cape Town, South Africa. *BMC Infectious Diseases* 16(1): 1–12. https://doi.org/10.1186/s12879-016-1934-z.

Dube, F.S., Ramjith, J., Gardner-Lubbe, S., Nduru, P., Robberts, F.J.L., Wolter, N., Zar, H.J., and Nicol, M.P. 2018. Longitudinal characterization of nasopharyngeal colonization with *Streptococcus pneumoniae* in a South African birth cohort post 13-valent pneumococcal conjugate vaccine implementation. *Scientific Reports* 8(1): 1–9. https://doi.org/10.1038/s41598-018-30345-5.

Ekoru, K., Young, E.H., Adebamowo, C., Balde, N., Hennig, B.J., Kaleebu, P., Kapiga, S., et al. 2016. H3Africa multi-centre study of the prevalence and environmental and genetic determinants of type 2 diabetes in sub-Saharan Africa: Study protocol. *Global Health, Epidemiology and Genomics*. 1(e5): 1–12. https://doi.org/10.1017/gheg.2015.6.

Fadlelmola, F.M., Panji, S., Ahmed, A.E., Ghouila, A., Akurugu, W.A., Domelevo Entfellner, J.B., Souiai, O., et al. 2019. Ten simple rules for organizing a webinar series. *PLoS Computational Biology*. 15(4): 1–7. https://doi.org/10.1371/journal.pcbi.1006671.

Feigin, V.L., Roth, G.A., Naghavi, M., Parmar, P., Krishnamurthi, R., Chugh, S., Mensah, G.A., et al. 2016. Global burden of stroke and risk factors in 188 countries, during 1990–2013: a systematic analysis for the Global Burden of Disease study 2013. *The Lancet Neurology* 15(9): 913–924. https://doi.org/10.1016/S1474-4422(16)30073-4.

Folarin, O.A., Ehichioya, D., Schaffiner, S.F., Winnicki, S.M., Wohl, S., Eromon, P., West, K.L., et al. 2016. Ebola virus epidemiology and evolution in Nigeria. *Journal of Infectious Diseases* 214(October): S102–S109. https://doi.org/10.1093/infdis/jiw190.

Fonseca, V., Libin, P.J.K., Theys, K., Faria, N.R., Nunes, M.R.T., Restovic, M.I., Freire, M., et al. 2019. A computational method for the identification of Dengue, Zika and Chikungunya virus species and genotypes. *PLoS Neglected Tropical Diseases* 13(5): 1–15. https://doi.org/10.1371/journal.pntd.0007231.

Gabriel, A., and Przybylski, J. 2010. *Sickle-Cell Anemia: A Look at Global Haplotype Distribution*. Nature Education. Rutgers University: New Jersey, USA.

Gardiner, S.A., Laing, N., Mall, S., and Wonkam, A. 2019. Perceptions of parents of children with hearing loss of genetic origin in South Africa. *Journal of Community Genetics* 10(3): 325–333. https://doi.org/10.1007/s12687-018-0396-y.

Geard, A., Pule, G.D., Chetcha Chemegni, B., Ngo Bitoungui, V.J., Kengne, A.P., Chimusa, E.R., and Wonkam, A. 2017. Clinical and genetic predictors of renal dysfunctions in sickle cell anaemia in Cameroon. *British Journal of Haematology* 178(4): 629–639. https://doi.org/10.1111/bjh.14724.

George J.A., Brandenburg J.T., Fabian J., Crowther N.J., Agongo G., Alberts M., Ali S., Asiki G., Boua P.R., Gómez-Olivé F.X., Mashinya F., Micksfield L., Mohamed S.F.,
Mukomana F., Norris S.A., Oduro A.R., Soo C., Sorgho H., Wade A., Naicker S., Ramsay M. 2019 AWI-Gen and the H3Africa Consortium. Kidney damage and associated risk factors in rural and urban sub-Saharan Africa (AWI-Gen): a cross-sectional population study. Lancet Global Health 7(12): e1632-e1643. https://doi.org/10.1016/S2214-109X(19)30443-7.

Geza, E., Mugo, J., Mulder, N.J., Wonkam, A., Chimusa, E.R., and Mazandu, G.K. 2019. A comprehensive survey of models for dissecting local ancestry deconvolution in human genome. Briefings in Bioinformatics 20(5): 1709–1724 https://doi.org/10.1093/bib/bby044.

Ghouila, A., Siwo, G.H., Entfellner, J.B.D., Panji, S., Button-Simons, K.A., Davis, S.Z., Fadlelmola, F.M., et al. 2018. Hackathons as a means of accelerating scientific discoveries and knowledge transfer. Genome Research 28(5): 759–765. https://doi.org/10.1101/gr.228460.117.

Gómez-Olivé, F.X., Ali, S.A., Made, F., Kyobutungi, C., Nonterah, E., Micklesfield, L., Alberts, M., et al. 2017. Regional and sex differences in the prevalence and awareness of hypertension: An H3Africa AWI-Gen Study across 6 sites in sub-Saharan Africa. Global Heart 12(2): 81–90. https://doi.org/10.1016/j.gheart.2017.01.007.

Guinto, C.O., Diarra, S., Diallo, S., Cissé, L., Coulibaly, T., Diallo, S.H., Tamégé, A., et al. 2017. A novel mutation in KIF5A in a Malian family with spastic paraplegia and sensory loss. Annals of Clinical and Translational Neurology 4(4): 272–275. https://doi.org/10.1002/acn3.402.

Gulsuner, S., Stein, D.J., Susser, E.S., Sibeko, G., Pretorius, A., Walsh, T., Majara, L., et al. 2020. Genetics of schizophrenia in the South African Xhosa. Science 367(6477): 569–573. https://doi.org/10.1126/science.aay8833.

Gurwitz, K.T., Aron, S., Panji, S., Maslamoney, S., Fernandes, P.L., Judge, D.P., Ghouila, A., et al. 2017. Designing a course model for distance-based online bioinformatics training in Africa: The H3ABioNet experience. PLoS Computational Biology 13(10): 1–11. https://doi.org/10.1371/journal.pcbi.1005715.

Haregu, T.N., Mohamed, S.F., Muthuri, S., Khayeka-Wandabwa, C., and Kyobutungi, C. 2018. Body mass index and wealth index: positively correlated indicators of health and wealth inequalities in Nairobi slums. Global Health, Epidemiology and Genomics 3(e11). https://doi.org/10.1017/gheg.2018.10.

Hendry, L.M., Sahibdeen, V., Choudhury, A., Norris, S.A., Ramsay, M., and Lombard, Z. 2018. Insights into the genetics of blood pressure in black South African individuals: The birth to twenty cohort. BMC Medical Genomics 11(1): 1–9. https://doi.org/10.1186/s12920-018-0321-6.

Hernández-de-Diego, R., Villiers, E.P. de, Klingström, T., Gourlé, H., Conesa, A., and Bongcam-Rudloff, E. 2017. The eBioKit, a stand-alone educational platform for bioinformatics. PLoS Computational Biology 13(9): 1–14. https://doi.org/10.1371/journal.pcbi.1005616.
Hotchkiss, J., Manyisa, N., Adadey, S.M., Oluwole, O.G., Wonkam, E., Mnika, K., Yalcouye, A., et al. 2019. The hearing impairment ontology: A tool for unifying hearing impairment knowledge to enhance collaborative research. *Genes* 21(10): 12. https://doi.org/10.3390/genes10120960.

Ilboudo, H., Noyes, H., Mulindwa, J., Kimuda, M.P., Koffi, M., Kaboré, J.W., Ahouty, B., et al. 2017. Introducing the TrypanoGEN biobank: A valuable resource for the elimination of human African trypanosomiasis. *PLoS Neglected Tropical Diseases* 11(6): e0005438. https://doi.org/10.1371/journal.pntd.0005438.

Jedy-Agba, E.E., Dareng, E.O., Adebamowo, S.N., Odutola, M., Oga, E.A., Igbinoba, F., Otu, T., et al. 2016. The burden of HPV associated cancers in two regions in Nigeria 2012–2014. *Cancer Epidemiology* 45(December): 91–97. https://doi.org/10.1016/j.canep.2016.10.008.

Jenkins, C., Arulogun, O.S., and Sarfo, S. 2017. Stroke investigative research and education network: Public outreach and engagement. *Journal of Community Medicine & Health Education* 7(2): 1–6. https://doi.org/10.4172/2161-0711.1000518.

Jenkins, C., Arulogun, O.S., Singh, A., Mande, A.T., Ajayi, E., Calys-Tagoe, B., Ovbiagele, B., et al. 2016. Stroke investigative research and education network: Community engagement and outreach within phenomics core. *Health Education & Behavior: The Official Publication of the Society for Public Health Education* 43(1 Suppl): 82S–92S. https://doi.org/10.1177/1090198116634082.

Jenkins, C., Burkett, N.-S., Ovbiagele, B., Mueller, M., Patel, S., Brunner-Jackson, B., Saulson, R., and Treiber, F. 2016. Stroke patients and their attitudes toward mHealth monitoring to support blood pressure control and medication adherence. *MHealth* 2(June): 24–24. https://doi.org/10.21037/mhealth.2016.05.04.

Jenkins, C., Ovbiagele, B., Arulogun, O., Singh, A., Calys-Tagoe, B., Akinyemi, R., Mande, A., et al. 2018. Knowledge, attitudes and practices related to stroke in Ghana and Nigeria: A SIREN call to action. *PLoS One* 13(11): 1–19. https://doi.org/10.1371/journal.pone.0206548.

Jongeneel, C.V., Achinike-Oduaran, O., Adebiyi, E., Adebiyi, M., Adeyemi, S., Akanle, B., Aron, S., et al. 2017. Assessing computational genomics skills: Our experience in the H3ABioNet African bioinformatics network. *PLoS Computational Biology* 13(6): 1–10. https://doi.org/10.1371/journal.pcbi.1005419.

Kaboré, J.W., Ilboudo, H., Noyes, H., Camara, O., Kaboré, J., Camara, M., Koffi, M., et al. 2017. Candidate gene polymorphisms study between human African trypanosomiasis clinical phenotypes in Guinea. *PLoS Neglected Tropical Diseases* 11(8): 1–13. https://doi.org/10.1371/journal.pntd.0005833.

Kamga, K.K., Nguefack, S., Minka, K., Tingang, E.W., Esterhuizen, A., Munung, S.N., Vries, J. de and Wonkam, A. 2020. Cascade testing for fragile X syndrome in a rural setting in Cameroon (Sub-Saharan Africa). *Genes* 11(2): 1–11. https://doi.org/10.3390/genes11020136.
Kamoto, K., Noyes, H., Nambala, P., Senga, E., Musaya, J., Kumwenda, B., Bucheton, B., et al. 2019. Association of APOL1 renal disease risk alleles with Trypanosoma brucei rhodesiense infection outcomes in the northern part of Malawi. *PLoS Neglected Tropical Diseases* 13(8): 1–12. https://doi.org/10.1371/journal.pntd.0007603.

Kimuda, M.P., Noyes, H., Mulindwa, J., Enyaru, J., Alibu, V.P., Sidibe, I., Mumba Ngoyi, D., et al. 2018. No evidence for association between APOL1 kidney disease risk alleles and human African trypanosomiasis in two Ugandan populations. *PLoS Neglected Tropical Diseases* 12(2): e0006300 https://doi.org/10.1371/journal.pntd.0006300.

Kumuthini, J., Zass, L., Panji, S., Salifu, S.P., Kayondo, J.K., Nembaware, V., Mbiyavanga, M., et al. 2019. The H3ABioNet helpdesk: An online bioinformatics resource, enhancing Africa’s capacity for genomics research. *BMC Bioinformatics* 20(1): 1–7. https://doi.org/10.1186/s12859-019-3322-3.

Landouré, G., Cissé, L., Touré, B.A., Yalcouyé, A., Coulibaly, T., Karambé, M., Sissoko, A.S., Coulibaly, T., Wonkam, A., and Guinto, C.O. 2017. Neurological complications in subjects with sickle cell disease or trait: Genetic results from Mali. *Global Heart* 12(2): 77–80. https://doi.org/10.1016/j.gheart.2017.01.014.

Landouré, G., Dembélé, K., Cissé, L., Samassékou, O., Diarra, S., Bocoum, A., Dembélé, M.E., Fischbeck, K.H., and Guinto, C.O. 2019. Hereditary spastic paraplegia type 35 in a family from Mali. *American Journal of Medical Genetics, Part A* 179(7): 1122–1125. https://doi.org/10.1002/ajmg.a.6179.

Landouré, G., Dembélé, K., Diarra, S., Cissé, L., Samassékou, O., Bocoum, A., Yalcouyé, A., Traoré, M., Fischbeck, K.H., and Guinto, C.O. 2020. A novel variant in the spastacin gene causing SPG11 in a Malian family. *Journal of the Neurological Sciences*. 411(November): 116675. https://doi.org/10.1016/j.jns.2020.116675.

Makubi, A., Mmbando, B.P., Novelli, E.M., Lwakatare, J., Soka, D., Marik, H., Tibarazwa, K., et al. 2017. Rates and risk factors of hypertension in adolescents and adults with sickle cell anaemia in Tanzania: 10 years’ experience. *British Journal of Haematology* 177(6): 930–937. https://doi.org/10.1111/bjh.14330.

Manenzhe, R.I., Moodley, C., Abdulgader, S.M., Robberts, F.J.L., Zar, H.J., Nicol, M.P., and Dube, F.S. 2019. Nasopharyngeal carriage of antimicrobial-resistant pneumococci in an intensively sampled South African Birth Cohort. *Frontiers in Microbiology* 10(March): 1–10. https://doi.org/10.3389/fmicb.2019.00610.

Mashinya, F., Alberts, M., Cook, I., and Ntuli, S. 2018. Determinants of body mass index by gender in the Dikgale Health and Demographic Surveillance System site, South Africa. *Global Health Action* 11(sup2): 1–12. https://doi.org/10.1080/16549716.2018.1537613.

Masiye, F., Mayosi, B., and Vries, J. de. 2017. “I passed the test!” Evidence of diagnostic misconception in the recruitment of population controls for an H3Africa genomic study in Cape Town, South Africa. *BMC Medical Ethics* 18(1): 1–9. https://doi.org/10.1186/s12910-017-0175-z.
Matimba, A., Tybring, G., Chitereka, J., Zinyama-Gutsire, R., Dandara, C., Bürén, E., Dhoro, M., and Masimirembwa, C. 2016. Practical Approach to Biobanking in Zimbabwe: Establishment of an Inclusive Stakeholder Framework. Biopreservation and Biobanking 14(5): 440–446. https://doi.org/10.1089/bio.2015.0043.

Mboowa, G., Mwesigwa, S., Katagirya, E., Retshabile, G., Mlotshwa, B.C., Williams, L., Kekitiinwa, A., et al. 2018. 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 Research 1(May): 1–16. https://doi.org/10.12688/aasopenres.12832.1.

Micklesfield, L.K., Kagura, J., Munthali, R., Crowther, N.J., Jaff, N., Gradidge, P., Ramsay, M., and Norris, S.A. 2018. Demographic, socio-economic and behavioural correlates of BMI in middle-aged black men and women from urban Johannesburg, South Africa. Global Health Action 11(sup2): 1–13. https://doi.org/10.1080/16549716.2018.148250.

Mnika, K., Mazandu, G.K., Jonas, M., Pule, G.D., Chimusa, E.R., Hanchard, N.A., and Wonkam, A. 2019. Hydroxyurea-induced miRNA expression in sickle cell disease patients in Africa. Frontiers in Genetics 10(MAY): 1–6. https://doi.org/10.3389/fgene.2019.00509.

Mohamed, S.F., Haregu, T.N., Khayeka-Wandabwa, C., Muthuri, S.K., and Kyobutungi, C. 2019. Magnitude and predictors of normal-weight central obesity— the AWI-Gen study findings. Global Health Action 12(1): 1–10. https://doi.org/10.1080/16549716.2019.1685809.

Moodley, K., and Beyer, C. 2019. Tygerberg research Ubuntu-inspired community engagement model: Integrating community engagement into genomic biobanking. Biopreservation and Biobanking. 17(6): 613–624. https://doi.org/10.1089/bio.2018.0136.

Moodley, K., and Singh, S. 2016. “It's all about trust”: reflections of researchers on the complexity and controversy surrounding biobanking in South Africa. BMC Medical Ethics 17(1): 1–9. https://doi.org/10.1186/s12910-016-0140-2.

Mulder, N., Abimiku, A., Adebamowo, S.N., Vries, J. de, Matimba, A., Olowoyo, P., Ramsay, M., Skelton, M., and Stein, D.J. 2018. H3Africa: current perspectives. Pharmacogenomics and Personalized Medicine 11(April): 59–66. https://doi.org/10.2147/PGPM.S141546.

Mulder, N., Nembaware, V., Adekile, A., Anie, K.A., Inusa, B., Brown, B., Campbell, A., et al. 2016. Proceedings of a sickle cell disease ontology workshop –Towards the first comprehensive ontology for sickle cell disease. Applied and Translational Genomics 9(March): 23–29. https://doi.org/10.1016/j.atg.2016.03.005.

Mulder, N.J., Adebiyi, E., Adebiyi, M., Adedayemi, S., Ahmed, A., Ahmed, R., Akanle, B., et al. 2017. Development of bioinformatics infrastructure for genomics research. Global Heart 12(2): 91–98. https://doi.org/10.1016/j.gh.heart.2017.01.005.
Mulder, N.J., Adebiyi, E., Alami, R., Benkahla, A., Brandful, J., Doumbia, S., Everett, D., et al. 2016. H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. Genome Research 26(2): 271–277. https://doi.org/10.1101/gr.196295.115.

Mulindwa, J., Noyes, H., Ilboudo, H., Pagani, L., Nyangiri, O., Kimuda, M.P., Ahouty, B., et al. 2020. High levels of genetic diversity within nilo-saharan populations: Implications for human adaptation. American Journal of Human Genetics 107(3): 473–486. https://doi.org/10.1016/j.ajhg.2020.07.007.

Munung, N.S., Marshall, P., Campbell, M., Littler, K., Masiye, F., Ouwe-Missi-Oukem-Boy, O., Seeley, J., Stein, D.J., Tindana, P., and Vries, J. de. 2016. Obtaining informed consent for genomics research in Africa: Analysis of H3Africa consent documents. Journal of Medical Ethics 42(2): 132–137. https://doi.org/10.1136/medethics-2015-102796.

Munung, N.S., Mayosi, B.M., and Vries, J. de. 2017. Equity in international health research collaborations in Africa: Perceptions and expectations of African researchers. PLoS One 12(10): 1–17. https://doi.org/10.1371/journal.pone.0186237.

Nembaware, V., and Mulder, N. 2019. The African Genomic Medicine Training Initiative (AGMT): Showcasing a community and framework driven genomic medicine training for nurses in Africa. Frontiers in Genetics 10(December): 1–10. https://doi.org/10.3389/fgene.2019.01209.

Nonterah, E.A., Boua, P.R., Klipstein-Grobusch, K., Asiki, G., Micklesfield, L.K., Agongo, G., Ali, S.A., et al. 2019. Classical cardiovascular risk factors and HIV are associated with carotid intima-media thickness in adults from sub-Saharan Africa: Findings from H3Africa AWI-Gen Study. Journal of the American Heart Association 8(14): e011506. https://doi.org/10.1161/JAHA.118.011506.

Nonterah, E.A., Debpuur, C., Agongo, G., Amenga-Etego, L., Crowther, N.J., Ramsay, M., and Rexford Oduro, A. 2018. Socio-demographic and behavioural determinants of body mass index among an adult population in rural Northern Ghana: the AWI-Gen study. Global Health Action 11(November): 1–11. https://doi.org/10.1080/16549716.2018.1467588.

Nuru, A., Mamo, G., Worku, A., Admasu, A., Medhin, G., Pieper, R., and Ameni, G. 2015. Genetic Diversity of Mycobacterium tuberculosis complex isolated from tuberculosis patients in Bahir Dar city and its surroundings, northwest Ethiopia. BioMed Research International 2015. https://doi.org/10.1155/2015/174732.

Odutola, M., Jedy-Agba, E.E., Dareng, E.O., Oga, E.A., Igbinoba, F., Otu, T., Ezeome, E., Hassan, R., and Adebamowo, C.A. 2016. Burden of cancers attributable to infectious agents in Nigeria: 2012–2014. Frontiers in Oncology 6(OCT): 1–17. https://doi.org/10.3389/fonc.2016.00216.

Ofon, E., Noyes, H., Ebo’o Eyanga, V., Njiokou, F., Koffi, M., Fogue, P., Hertz-Fowler, C., MacLeod, A., Matovu, E., and Simo, G. 2018. Association between IL1 gene polymor-
phism and human African trypanosomiasis in populations of sleeping sickness foci of southern Cameroon. *PLoS Neglected Tropical Diseases* 13(3): 1–21. https://doi.org/10.1371/journal.pntd.0007283.

Ofon, E., Noyes, H., Mulindwa, J., Ilboudo, H., Simuunza, M., Ebo'o, V., Njiokou, F., et al. 2017. A polymorphism in the haptoglobin, haptoglobin related protein locus is associated with risk of human sleeping sickness within Cameroonian populations. *PLoS Neglected Tropical Diseases* 11(10): 1–16. https://doi.org/10.1371/journal.pntd.0005979.

Ojagbemi, A., Owolabi, M., Akinyemi, R., Arulogun, O., Akinyemi, J., Akpa, O., Sarfo, F.S., et al. 2017. Prevalence and predictors of anxiety in an African sample of recent stroke survivors. *Acta Neurologica Scandinavica* 136(6): 617–623. https://doi.org/10.1111/ane.12766.

Ojagbemi, Akin, Owolabi, M., Akinyemi, J., and Ovbiagele, B. 2017a. Criterion validity of the “HRQOLISP-E”: A new context-specific screening tool for poststroke depression. *Behavioural Neurology* 2017. https://doi.org/10.1155/2017/6515769.

Ojagbemi, Akin, Owolabi, M., Akinyemi, J., and Ovbiagele, B. 2017b. Proposing a new stroke-specific screening tool for depression: Examination of construct validity and reliability. *ENeurologicalSci* 9(December): 14–18. https://doi.org/10.1016/j.ensci.2017.10.002.

Olowoyo, P., Owolabi, M.O., Fawale, B., and Ogguniyi, A. 2016. Short term stroke outcome is worse among individuals with sickle cell trait. *ENeurologicalSci* 3(June): 64–68. https://doi.org/10.1016/j.ensci.2016.02.009.

Osafo, C., Raji, Y.R., Burke, D., Tayo, B.O., Tiffin, N., Moxey-Mims, M.M., Rasooly, R.S., et al. 2015. Human Heredity and Health (H3) in Africa kidney disease research network: A focus on methods in sub-Saharan Africa. *Clinical Journal of the American Society of Nephrology* 10(12): 2279–2287. https://doi.org/10.2215/CJN.11951214.

Osafo, C., Raji, Y.R., Olanrewaju, T., Mamven, M., Arogundade, F., Ajayi, S., Ulasi, I., et al. 2016. Genomic approaches to the burden of kidney disease in Sub-Saharan Africa: the Human Heredity and Health in Africa (H3Africa) Kidney Disease Research Network. *Kidney International* 90(1): 2–5. https://doi.org/10.1016/j.kint.2015.12.059.

Owolabi, M. O., Akpa, O.M., and Agunloye, A.M. 2016. Carotid IMT is more associated with stroke than risk calculators. *Acta Neurologica Scandinavica* 133(6): 442–450. https://doi.org/10.1111/ane.12482.

Owolabi, M., Sarfo, F., Howard, V.J., Irvin, M.R., Gebregziabher, M., Akinyemi, R., Bennett, A., et al. 2017. Stroke in indigenous Africans, African Americans, and European Americans: Interplay of racial and geographic factors. *Stroke* 48(5): 1169–1175. https://doi.org/10.1161/STROKEAHA.116.015937.

Owolabi, Mayowa O, Akpa, O.M., Made, F., Adebamowo, S.N., Ojo, A., Adu, D., Motala, A.A., et al. 2019. Data resource profile: Cardiovascular H3Africa Innovation
Resource (CHAIR). *International Journal of Epidemiology* 48(2): 366–367. https://doi.org/10.1093/ije/dyy261.

Owolabi, Mayowa Ojo, Sarfo, F., Akinyemi, R., Gebregziabher, M., Akpa, O., Akpalu, A., Wahab, K., et al. 2018. Dominant modifiable risk factors for stroke in Ghana and Nigeria (SIREN): a case-control study. *The Lancet Global Health* 6(4): e436–e446. https://doi.org/10.1016/S2214-109X(18)30002-9.

Pisa, P.T., Micklesfield, L.K., Kagura, J., Ramsay, M., Crowther, N.J., and Norris, S.A. 2018. Different adiposity indices and their association with blood pressure and hypertension in middle-aged urban black South African men and women: Findings from the AWI-GEN South African Soweto Site. *BMC Public Health* 18(1): 1–8. https://doi.org/10.1186/s12889-018-5443-4.

Pratt, B., and Vries, J. de. 2018. Community engagement in global health research that advances health equity. *Bioethics* 32(7): 454–463. https://doi.org/10.1111/bioe.12465.

Pule, G.D., Bitoungui, V.J.N., Chemegni, B.C., Kengne, A.P., and Wonkam, A. 2017. SARIa promoter polymorphisms are not associated with fetal hemoglobin in patients with sickle cell disease from Cameroon. *BMC Research Notes* 10(1): 1–5. https://doi.org/10.1186/s13104-017-2502-3.

Pulis, G.D., Ngo Bitoungui, V.J., Chetcha Chemegni, B., Kengne, A.P., Antonarakis, S., and Wonkam, A. 2015. Association between variants at BCL11A erythroid-specific enhancer and fetal hemoglobin levels among sickle cell disease patients in Cameroon: Implications for future therapeutic interventions. *Omics: A Journal of Integrative Biology* 19(10): 627–631. https://doi.org/10.1089/omi.2015.0124.

Pule, G.Dineo., Mnika, K., Joubert, M., Mowla, S., Novitzky, N., and Wonkam, A. 2017. Burden, genotype and phenotype profiles of adult patients with sickle cell disease in Cape Town, South Africa. *South African Medical Journal* 107(2): 149–155. https://doi.org/10.7196/SAMJ.2017.v107i2.10849.

Ramsay, M., Crowther, N., Tambo, E., Agongo, G., Baloyi, V., Dikotope, S., Gómez-Olivé, X., et al. 2016. H3Africa AWI-Gen Collaborative Centre: A resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. *Global Health, Epidemiology and Genomics* 1(e20): 1–13. https://doi.org/10.1017/gheg.2016.17.

Ramsay, M., Crowther, N.J., Agongo, G., Ali, S.A., Asiki, G., Boua, R.P., Gómez-Olivé, F.X., et al. 2018. Regional and sex-specific variation in BMI distribution in four sub-Saharan African countries: The H3Africa AWI-Gen study. *Global Health Action* 11(sup2): 90–97. https://doi.org/10.1080/16549716.2018.1556561.

Ramsay, M., Vries, J. de, Soodyall, H., Norris, S.A., and Sankoh, O. 2014. Ethical issues in genomic research on the African continent: Experiences and challenges to ethics review committees. *Human Genomics*. 8(1): 4–9 https://doi.org/10.1186/s40246-014-0015-x.
Retshabile, G., Mlotshwa, B.C., Williams, L., Mwesigwa, S., Mboowa, G., Huang, Z., Rustagi, N., et al. 2018. Whole-exome sequencing reveals uncaptured variation and distinct ancestry in the Southern African population of Botswana. *American Journal of Human Genetics* 102(5): 731–743. https://doi.org/10.1016/j.ajhg.2018.03.010.

Ross, C., Nizami, B., Glenister, M., Amanu, O.S., Atigheh, A.R., Atigheh, C., and Bishop, Ö.T. 2018. MODE-TASK: Large-scale protein motion tools. *Bioinformatics* 34(21): 3759–3763. https://doi.org/10.1093/bioinformatics/bty427.

Rotimi, C., Aban, A., Abimiku, A., Adabayeri, V.M., Adabamowo, C., Adebiyi, E., Ademola, A.D., et al. 2014. Research capacity. Enabling the genomic revolution in Africa. *Science*. 344(6190): 1346–1348. https://doi.org/10.1126/science.1251546.

Rudan, I., O’Brien, K.L., Nair, H., Liu, L., Theodoratou, E., Qazi, S., Lukić, I., et al. 2013. Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. *Journal of Global Health* 3(1): 10401. https://doi.org/10.7189/jogh.03.010401.

Sahibdeen, V., Crowther, N.J., Soodyall, H., Hendry, L.M., Munthali, R.J., Hazellhurst, S., Choudhury, A., Norris, S.A., Ramsay, M., and Lombard, Z. 2018. Genetic variants in SEC61B are associated with body composition in black South Africans. *Nutrition and Diabetes* 8(1): 1–13. https://doi.org/10.1038/s41387-018-0050-0.

Sarfo, F., Gebregziabher, M., Owbiagele, B., Akinyemi, R., Owolabi, L., Obiako, R., Akpa, O., et al. 2016. Multilingual validation of the questionnaire for verifying stroke-free status in West Africa. *Stroke* 47(1): 167–172. https://doi.org/10.1161/STROKEAHA.115.010374.

Sarfo, F. S., Gebregziabher, M., Owbiagele, B., Akinyemi, R., Owolabi, L., Obiako, R., Armstrong, K., et al. 2016. Validation of the 8-item questionnaire for verifying stroke free status with and without pictograms in three West African languages. *ENeurologicalSci* 3(June): 75–79. https://doi.org/10.1016/j.ensci.2016.03.004.

Sarfo, F. S., Gyamfi, R.A., Adamu, S., Sarfo-Kantanka, O., Owolabi, M., and Owbiagele, B. 2017. Administration of a pictorial questionnaire to screen for stroke among patients with hypertension or diabetes in rural Ghana. *Journal of the Neurological Sciences* 373(Febuary): 289–294. https://doi.org/10.1016/j.jns.2017.01.022.

Sarfo, F. S., Nichols, M., Qanungo, S., Teklehaimanot, A., Singh, A., Mensah, N., Saulson, R., et al. 2017. Stroke-related stigma among West Africans: Patterns and predictors. *Journal of the Neurological Sciences* 375(April): 270–274. https://doi.org/10.1016/j.jns.2017.02.018.

Sarfo, F. S., Opare-Sem, O., Agyei, M., Akassi, J., Owusu, D., Owolabi, M., and Owbiagele, B. 2018. Risk factors for stroke occurrence in a low HIV endemic West African country: A case-control study. *Journal of the Neurological Sciences* 395(December): 8–16. https://doi.org/10.1016/j.jns.2018.09.021.
Sarfo, F. S., Ovbiagele, B., Akassi, J., and Kyem, G. 2017. Baseline prescription and one-year persistence of secondary prevention drugs after an index stroke in central Ghana. *ENeurologicalSci* 6(March): 68–73. https://doi.org/10.1016/j.ensci.2016.12.003.

Sarfo, F. S., Ovbiagele, B., Gebregziabher, M., Wahab, K., Akinyemi, R., Akpalu, A., Akpa, O., et al. 2018. Stroke among young West Africans: Evidence from the SIREN (stroke investigative research and educational network) large multisite case-control study. *Stroke* 49(5): 1116–1120. https://doi.org/10.1161/STROKEAHA.118.020783.

Sengupta, D., Choudhury, A., Fortes-Lima, C., Aron, S., Whitelaw, G., Bostoen, K., Gunnink, H., et al. 2021. Genetic substructure and complex demographic history of South African Bantu speakers. *Nature Communications* 12(1): 2080. https://doi.org/10.1038/s41467-021-22207-y.

Shaffer, J.G., Mather, F.J., Wele, M., Li, J., Tangara, C.O., Kassogue, Y., Srivastav, S.K., et al. 2019. Expanding research capacity in sub-Saharan Africa through informatics, bioinformatics, and data science training programs in Mali. *Frontiers in Genetics* 10(APR): 1–13. https://doi.org/10.3389/fgene.2019.00331.

Siddle, K.J., Eromon, P., Barnes, K.G., Mehta, S., Ogunie, J.U., Odia, I., Schaffner, S.F., et al. 2018. Genomic analysis of Lassa virus during an increase in cases in Nigeria in 2018. *The New England Journal of Medicine* 379(18): 1745–1753. https://doi.org/10.1056/NEJMoa1804498.

Singh, A., Jenkins, C., Calys-Tagoe, B., Arulogun, O.S., Sarfo, S., Ovbiagele, B., Akpalu, A., Melikam, S., Uvere, E., and Owolabi, M.O. 2017. Stroke investigative research and education network: Public outreach and engagement. *Journal of Community Medicine & Health Education* 7(2): 518. https://doi.org/10.4172/2161-0711.1000518.

Soo, C.C., Mukomana, F., Hazelhurst, S., and Ramsay, M. 2017. Establishing an academic biobank in a resource-challenged environment. *South African Medical Journal* 107(6): 486–492. https://doi.org/10.7196/SAMJ.2017.v107i6.12099.

Sood, S., Ojo, A.O., Adu, D., Kannan, K., Ghassabian, A., Koshy, T., Vento, S.M., et al. 2019. Association between perfluoroalkyl substance exposure and renal function in children with CKD enrolled in H3Africa Kidney Disease Research Network. *Kidney International Reports* 4(11): 1641–1645. https://doi.org/10.1016/j.eikir.2019.07.017.

Staunton, C., and Moodley, K. 2016. Data mining and biological sample exportation from South Africa: A new wave of bioexploitation under the guise of clinical care? *SAMJ: South African Medical Journal* 106: 136–138. https://doi.org/10.7196/SAMJ.2016.v106i2.10248

Staunton, Ciara, Abayomi, A., Bassa, F., and Moodley, K. 2019. Negotiating requests for reimbursement for community engagement: Challenges in developing an educational video for Genomic Biobanking Research in South Africa. *Journal of Empirical Research on Human Research Ethics* 14(5): 501–503. https://doi.org/10.1177/1556264619856223.
Staunton, Ciara and Moodley, K. 2013. Challenges in biobank governance in Sub-Saharan Africa. *BMC Medical Ethics* 14(1): 355. https://doi.org/10.1186/1472-6939-14-35.

Staunton, Ciara, Tindana, P., Hendricks, M., and Moodley, K. 2018. Rules of engagement: Perspectives on stakeholder engagement for genomic biobanking research in South Africa. *BMC Medical Ethics* 19(1): 1–10. https://doi.org/10.1186/s12910-018-0252-y.

Stremlau, M.H., Andersen, K.G., Folarin, O.A., Grove, J.N., Odia, I., Ehiane, P.E., Omoniwa, O., et al. 2015. Discovery of novel rhabdoviruses in the blood of healthy individuals from West Africa. *PLoS Neglected Tropical Diseases* 9(3): 1–17. https://doi.org/10.1371/journal.pntd.0003631.

Tekendo-Ngongang, C., Dahoun, S., Nguefack, S., Moix, I., Gimelli, S., Zambo, H., Morris, M.A., Sloan-Béna, F., and Wonkam, A. 2020. MECP2 duplication syndrome in a patient from Cameroon. *American Journal of Medical Genetics, Part A* 2( November): 1–4. https://doi.org/10.1002/ajmg.a.61510.

Teklu, T., Kwon, K., Wondale, B., HaileMariam, M., Zewude, A., Medhin, G., Legesse, M., Pieper, R., and Ameni, G. 2018. Potential immunological biomarkers for detection of *Mycobacterium tuberculosis* infection in a setting where *M. tuberculosis* is endemic, Ethiopia. *Infection and Immunity* 86(4): 1–11. https://doi.org/10.1128/IAI.00759-17.

Tiffin, N. 2018. Tiered informed consent: Respecting autonomy, agency and individuality in Africa. *BMJ Global Health* 6(3): e001249. https://doi.org/10.1136/bmjgh-2018-001249.

Tiffin, N., George, A., and Lefevre, A.E. 2019. How to use relevant data for maximal benefit with minimal risk: Digital health data governance to protect vulnerable populations in low-income and middle-income countries. *BMJ Global Health* 4(2): 1–9. https://doi.org/10.1136/bmjgh-2019-001395.

Tindana, P., Campbell, M., Marshall, P., Littler, K., Vincent, R., Seeley, J., Vries, J. de and Kamuya, D. 2017. Developing the science and methods of community engagement for genomic research and biobanking in Africa. *Global Health, Epidemiology and Genomics* 2(e13). https://doi.org/10.1017/gheg.2017.9.

Tindana, Paulina, Yakubu, A., Staunton, C., Matimba, A., Littler, K., Madden, E., Munung, N.S., and Vries, J. de. 2019. Engaging research ethics committees to develop an ethics and governance framework for best practices in genomic research and biobanking in Africa: The H3Africa model. *BMC Medical Ethics* 20(1): 1–7. https://doi.org/10.1186/s12910-019-0398-2.

Vanker, A., Nduru, P.M., Barnett, W., Dube, F.S., Gie, R.P., Nicol, M.P., and Zar, H.J. 2019. Indoor air pollution and tobacco smoke exposure: impact on nasopharyngeal bacterial carriage in mothers and infants in an African birth cohort study. *ERJ Open Research* 5(1): 00052–02018. https://doi.org/10.1183/23120541.00052-2018.
Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., Pauwels, R., et al. 2019. Genome Detective: An automated system for virus identification from high-throughput sequencing data. *Bioinformatics* 35(5): 871–873. https://doi.org/10.1093/bioinformatics/bty695.

Voight, B.F., Kang, H.M., Ding, J., Palmer, C.D., Sidore, C., Chines, P.S., Burtt, N.P., et al. 2012. The Metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genetics* 8(8): e1002793. https://doi.org/10.1371/journal.pgen.1002793.

Vries, J. de, Littler, K., Matimba, A., McCurdy, S., Ouwe Missi Oukem-Boyer, O., Seeley, J., and Tindana, P. 2016. Evolving perspectives on broad consent for genomics research and biobanking in Africa. Report of the second H3Africa Ethics Consultation Meeting, 11 May 2015. *Global Health, Epidemiology and Genomics* 1: e13. https://doi.org/10.1017/gheg.2016.5.

Vries, Jantina de, Abayomi, A., Brandful, J., Littler, K., Madden, E., Marshall, P., Ouwe Missi Oukem-Boyer, O., and Seeley, J. 2014. A perpetual source of DNA or something really different: Ethical issues in the creation of cell lines for African genomics research. *BMC Medical Ethics*. 15(1): 1–7. https://doi.org/10.1186/1472-6939-15-60.

Vries, Jantina de, Munung, S.N., Matimba, A., McCurdy, S., Ouwe Missi Oukem-Boyer, O., Staunton, C., Yakubu, A., and Tindana, P. 2017. Regulation of genomic and biobanking research in Africa: A content analysis of ethics guidelines, policies and procedures from 22 African countries. *BMC Medical Ethics* 18(1): 1–9. https://doi.org/10.1186/s12910-016-0165-6.

Wagner, R.G., Crowther, N.J., Gómez-Olivé, F.X., Kabudula, C., Kahn, K., Mhembere, M., Myakayaka, Z., Tollman, S., and Wade, A.N. 2018. Sociodemographic, socioeconomic, clinical and behavioural predictors of body mass index vary by sex in rural South African adults-findings from the AWI-Gen study. *Global Health Action* 11(sup2): 1549436. https://doi.org/10.1080/16549716.2018.1549436.

Wondale, B., Medhin, G., Abebe, G., Tolosa, S., Mohammed, T., Teklu, T., Pieper, R., and Ameni, G. 2018. Phenotypic and genotypic drug sensitivity of *Mycobacterium tuberculosis* complex isolated from south Omo zone, southern Ethiopia. *Infection and Drug Resistance* 11: 1581–1589. https://doi.org/10.2147/IDR.S165088.

Wonkam, A., and Hurst, S. 2014. A call for policy action in sub-saharan Africa to rethink diagnostics for pregnancy affected by sickle cell disease: Differential views of medical doctors, parents and adult patients predict value conflicts in cameroon. *OMICS A Journal of Integrative Biology* 18(7): 472–480. https://doi.org/10.1089/omi.2013.0167.

Wonkam, A., Mnika, K., Ngo Bitoungui, V.J., Chetcha Chemegni, B., Chimusa, E.R., Dandara, C., and Kengne, A.P. 2018. Clinical and genetic factors are associated with pain and hospitalisation rates in sickle cell anaemia in Cameroon. *British Journal of Haematology* 180(1): 134–146. https://doi.org/10.1111/bjh.15011.
Wonkam, A., and Vries, J. de. 2020. Returning incidental findings in African genomics research. *Nature Genetics* 52(1): 17–20. https://doi.org/10.1038/s41588-019-0542-4.

Wonkam, E.T., Chimusa, E., Noubiap, J.J., Adadey, S.M., Fokouo, J.V.F., and Wonkam, A. 2019. GJB2 and GJB6 mutations in hereditary recessive non-syndromic hearing impairment in Cameroon. *Genes* 10(11): 844. https://doi.org/10.3390/genes10110844.

Wright, G.E.B., Adeyemo, A.A., and Tiffin, N. 2014. Informed consent and ethical re-use of African genomic data. *Human Genomics* 8(1): 1–3. https://doi.org/10.1186/s40246-014-0018-7.

Yakubu, A., Tindana, P., Matimba, A., Littler, K., Munung, N.S., Madden, E., Staunton, C., and Vries, J. de. 2018. Model framework for governance of genomic research and biobanking in Africa – a content description. *AAS Open Research* 1(April): 13. https://doi.org/10.12688/aasopenres.12844.1.

Yalcouyé, A., Diallo, S.H., Coulibaly, T., Cissé, L., Diallo, S., Samassékou, O., Diarra, S., et al. 2019. A novel mutation in the GARS gene in a Malian family with Charcot-Marie-Tooth disease. *Molecular Genetics and Genomic Medicine* 7(7): 1–4. https://doi.org/10.1002/mgg3.782.