Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

The critical needs and challenges for genetic architecture studies in Africa

Alicia R Martin[1,2,3], Solomon Teferra[4,5], Marlo Möller[6], Eileen G Hoal[6] and Mark J Daly[1,2,3,7]

Human genetic studies have long been vastly Eurocentric, raising a key question about the generalizability of these study findings to other populations. Because humans originated in Africa, these populations retain more genetic diversity, and yet individuals of African descent have been tremendously underrepresented in genetic studies. The diversity in Africa affords ample opportunities to improve fine-mapping resolution for associated loci, discover novel genetic associations with phenotypes, build more generalizable genetic risk prediction models, and better understand the genetic architecture of complex traits and diseases subject to varying environmental pressures. Thus, it is both ethically and scientifically imperative that geneticists globally surmount challenges that have limited progress in African genetic studies to date. Additionally, African investigators need to be meaningfully included, as greater inclusivity and enhanced research capacity afford enormous opportunities to accelerate genomic discoveries that translate more effectively to all populations. We review the advantages, challenges, and examples of genetic architecture studies of complex traits and diseases in Africa. For example, with greater genetic diversity comes greater ancestral heterogeneity; this higher level of understudied diversity can yield novel genetic findings, but some methods that assume homogeneous population structure and work well in European populations may work less well in the presence of greater heterogeneity in African populations. Consequently, we advocate for methodological development that will accelerate studies important for all populations, especially those currently underrepresented in genetics.

Addresses
[1] Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA
[2] Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[3] Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[4] Department of Psychiatry, School of Medicine, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia
[5] Department of Epidemiology, Harvard T. H. Chan School of Public Health, Harvard University, Boston, USA
[6] DST-NRF Centre of Excellence for Biomedical Tuberculosis Research, South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, Cape Town, South Africa
[7] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

Corresponding author: Martin, Alicia R (armartin@broadinstitute.org)

Current Opinion in Genetics & Development 2018, **53**:113–120

This review comes from a themed issue on **Genetics of human origins**

Edited by **Lluis Quintana-Murci** and **Brenna Henn**

https://doi.org/10.1016/j.gde.2018.08.005

Historical biases in genetic studies

Nearly a decade ago, 96% of participants in genome-wide association studies (GWAS) were of European descent [1]. While European individuals now account for 78% of GWAS participants [2•], the non-European proportion has stagnated since 2014. African ancestry individuals constitute merely 2.4% of participants (although notably account for 7% of all associations) [2•]. This participant bias results in interpretability gaps by ancestry with medically relevant consequences [3,4]. For example, while easily avoidable, African American patients were more likely than white Americans to be incorrectly told they have a genetic mutation that increases their risk of hypertrophic cardiomyopathy, an early-onset life-threatening heart disease, at leading genetic testing labs [5]. Additionally, drug metabolism genes such as *CYP3A4* contain mutations that can alter dosage requirements, but pharmacogenetic variants are disproportionately uncatalogued among African populations [6], so genotype-based dosage guidelines are less useful. In the US, the National Human Genome Research Institute has prioritized increased diversity in genetic studies [7•]. This prioritization is an important step that, if heeded, will aid interpretations in medical genomics for all ethnicities [8]. Greater inclusivity of African populations in medical genomics is important for accelerating genomic discoveries, enabling reconstruction of modern human origins, producing results that can be translated across populations more accurately, identifying genetic associations with traits for variants absent elsewhere, and building research capacity in Africa.

Genetic study biases have not happened in a vacuum, but have had widespread consequences for GWAS tools and

resources in African populations. Genotyping arrays have traditionally been biased towards alleles most frequent and imputable in European populations [9,10], compounding biases in which GWAS identify variant associations most common in the study population [11•,12]. In contrast, array backbones prioritizing SNPs that maximally tag variants across all populations improve imputation performance, providing more even genomic coverage [13]. Perhaps more importantly, imputation panels are vastly Eurocentric, shortchanging representation of the greater haplotypic diversity present in Africans from deeper recombination history [12,14,15]. The most widely available African sequencing resources have biased representation towards African Americans and West Africans [8,12], leaving huge swaths of African diversity uncatalogued.

Existing challenges to surmount for African genetics studies

To empower African genetic studies and build capacity for research aiding biological understanding across a diverse swath of humanity, we review challenges that need to be confronted and continually addressed.

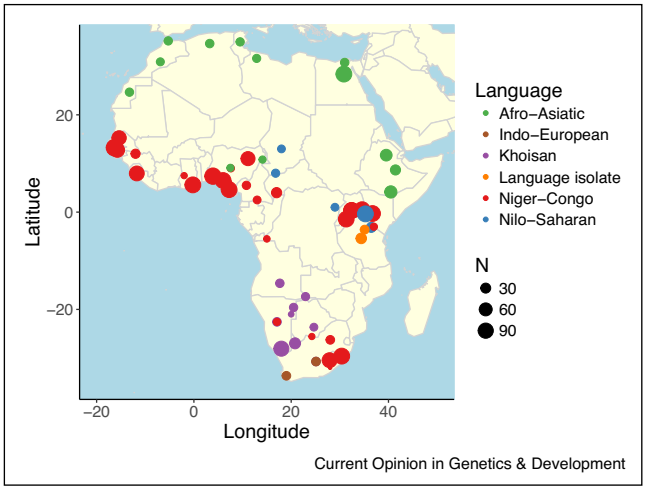
Historical

Africa has long been subjected to a violent and oppressive colonial history that has bred suspicion and an anticipation of resource exploitation. This understandable mistrust continues to strain ongoing relations, with new actors such as China in addition to European groups scrambling for African resources [16,17]. The impact on research collaborations is evident, with some authors discussing 'neo-colonial science' [18]. Such strained relations are more pronounced in collaborations involving genetic studies, especially when shipping samples out of Africa and the global south [19]. Some discuss 'genomic sovereignty' of Africans and ownership of African genetic material [20]. Proponents of international collaborations argue that working with high income countries will eventually ensure equity, justice, and benefit to Africans, with capacity building for genomic research providing immediate benefit for African institutions [21•], although concerns have been raised about the sustainability of these efforts. Ongoing tensions weigh the benefit to Africans by including more African researchers and DNA in global research against the challenges of promoting African science while integrating and importing the best science around the world into Africa (Figure 1).

Infrastructural

Conducting genetic studies in Africa is not an easy task. Infrastructural problems can include unreliable or no electricity in clinics and laboratories that process samples, impassable roads in some areas, and crime or political instability making some areas dangerous and/or inaccessible for researchers. Many African countries do not have sufficient laboratory equipment or facilities for genomics research, and most require imported reagents. Importing is not only time-consuming, but also costly—reagents are often many times more expensive in Africa than Western countries in real terms, not including shipment costs. Biobanks are less abundant, partially due to power interruptions affecting storage and processing of samples. Some African institutions have experience in large-scale human genetic analyses; the H3ABionet consortium has developed core bioinformatics infrastructure in Africa [22•]. However, high-speed internet connections and powerful computers are not always available to access large data files. Human resource issues can also be a challenge, namely high staff turnover due to inadequate pay, competing demands for time from qualified staff, and/or too few qualified staff. Relatedly, brain drain is a major issue, as many skilled African scientists leave the continent in search of greener pastures [23,24]. To be sensitive to these challenges, some major international research initiatives such as H3Africa have required a relatively long embargo period on publication for African researchers [25•]. Connecting African researchers to adequate computing power (e.g. stable wireless connections to cloud computing) may offer more direct means to facilitate research. Compared with the relative ease of acquiring samples in the global north, the focus of databanks on European/white populations is unsurprising, but it is nonetheless imperative that researchers rise to these challenges for the benefit of all.

Figure 1



Map of publicly available African samples and corresponding language families from previous studies. Reference data comes from several previous studies [12,15,38,39,83–87].

Funding

Genetics research is expensive, and a lack of attention from African policy makers in resource-limited settings is primarily driven by competing priorities for more immediate public health concerns, including infectious