

Opinion

African genomic data sharing and the struggle for equitable benefit

Michèle Ramsay^{1,*}¹Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa*Correspondence: michele.ramsay@wits.ac.za<https://doi.org/10.1016/j.patter.2021.100412>

Genomic and related health data from Africa remain scarce and are extremely valuable, due to an abundance of variants often rare or absent in the rest of the world. Insights from such data will benefit global populations, but will Africa be neglected by limited access to affordable benefits resulting from research that uses their data?

Africa is a treasure trove for genomic scientists worldwide. The reason for this is that genomes from Africans harbor unprecedented genetic diversity, and the potential for discovery is vast. For example, despite the steady increase in data in international repositories, high-depth whole-genome sequence data from just over 400 Africans yielded over 3 million novel, previously undocumented variants.¹ Understanding their potential health impact could benefit not just Africans, but the global community.

As a geneticist working in Africa, I find myself reflecting on the issue of data sharing. In a nutshell, there is great value in sharing data to enhance knowledge, but are we doing enough to ensure that people in Africa also benefit? How do we protect their interests when their data are used in research? We cannot accept a status quo that limits access to novel applied technologies in precision medicine only to those who can afford it, as the poor will continue to suffer and most of the poorest nations are in Africa.

Recent history documents research practices that have led to exploitation and misuse of research participant data and samples. Much has improved through considerable efforts of researchers, funders, and journal editorial processes, and the boxes are ticked with regard to ethics approval and consent. Unfortunately, meaningful engagement with researchers and communities in Africa leaves much to be desired. Nonetheless, my optimism rests on the emergence and development of more balanced and enabling research practices and on building trust to ensure that people in Africa have an opportunity to benefit

from research using their unique data resources.

African genetic and phenotype data are scarce and valuable

Genomic data from Africans, especially those with associated phenotype and health-relevant data, are scarce, extremely valuable, and highly desirable. This emphasizes our duty to the communities and research participants, as custodians of their data, to remain mindful of the need to carefully consider the nature of potential benefit. On a practical level, this could mean at least setting the scene for equitable and affordable access to translational science that makes it into a clinical setting.

Data sharing is essential to the advancement of science. How then do we balance access to data and data sharing with meaningful participation of African scientists in data analysis and knowledge generation? Working in a high-pressure environment where time and timely publication are of the essence could trump collegiality, erode trust, and adversely affect collaborations. Researchers in Africa often work under pressure of clinical and teaching commitments that dominate their working hours and have to navigate a lack of infrastructural support for managing general administrative tasks and procurement processes, less responsive ethics review committees, and little support for writing and managing research grants. Disparities in numbers of skilled researchers and the availability of resources, as well as the high cost of data collection, need to be considered. Systemic change is necessary to address these barriers, but there is no reason why collaborators and

funders cannot take these factors into consideration now when building partnerships and allocating budgets. The net effect would be ameliorating some of the inequities and promoting an enabling research culture that would also promote data sharing and the advancement of knowledge.

Positive interventions for a greater participation from African researchers include more balanced partnerships, capacity strengthening, and jointly developed global standards that are mindful of low-resource settings and cost-effective solutions. The Human Heredity and Health in Africa Consortium (<https://h3africa.org/>) has supported African-led genomic research, and the Global Alliance for Genomics and Health (GA4GH; <https://www.ga4gh.org/>) has developed standards toward promoting responsible international sharing of genomic data.

A salient aspect of research in Africa and for collaboration is the importance of working together to develop research questions that have impact and reflect the needs of African communities. African researchers need to take a central role to ensure that the way in which data are analyzed and the interpretation of the outcomes are taken into consideration and placed in the context of where and how the data were collected.

Too often Africa is represented by research in African diaspora communities in the US and Europe, but they should not be considered ideal proxies for Africa for three main reasons: (1) they generally include representation of limited geographic regions and ethnicities from Africa, (2) they live in different environments that can affect health outcomes through



gene-environment interactions, and (3) they are often admixed populations where individuals have varying degrees of African ancestry. It is therefore essential to study Africans who reside in different regions of the continent to more fully understand the impact of genetic variation on health.

How do we build a more sustainable culture that values research data in Africa?

Few would argue that health and genetic data from Africa have a finite lifetime in terms of relevance and potential for knowledge mining. Yet datasets at African institutions are often neglected once a project is concluded. They are either deleted or forgotten in an archive where their provenance is poorly documented, or reuse conditions are vague and uncertain, risking ethical breaches if reused.

Documenting research processes and developing governance frameworks, including informed consent processes, ethical oversight and reuse conditions for datasets, are essential elements that promote, build, and sustain community trust and cannot be overemphasized. African researchers have identified ethical principles and research practices to guide equity-oriented governance frameworks for health and genomic research that include reciprocity, open sharing, accountability, mutual trust, and inclusivity, as well as the need to level the playing field and promote fairness.²⁻⁴ These are essential elements to guard against potential exploitation and to promote fair practices and benefit sharing.

Worldwide researchers and institutions struggle to make data findable, accessible, interoperable, and reusable—meeting the FAIR principles. This struggle is exacerbated in Africa due to competing priorities, weak governance frameworks, and a lack of adequate resources. In the absence of efforts to reduce inequities and to recalibrate imbalances and with a lack of comparable protection mechanisms to guide data sharing and use, it is unlikely that potential benefits will be fairly distributed. These considerations need to be prioritized at African institutions and more broadly in African countries.

What are the challenges, and where is the benefit?

It is well recognized that it is more expensive to do research in Africa, and it takes

longer to complete projects. Equipment needs to be imported, and import taxes are high; maintenance and servicing of equipment costs more; expertise is limited; spare parts need to be sourced out of country; and there are long delays due to travel and availability of technicians. Dedicated research positions are rare, and the cycle of grant funding and project completion is relentless, meaning that data are usually not sufficiently mined and curated for future analyses. Funding for data storage, integration, and reuse is scarce or absent, and these activities are perceived to have low value, with funding for the generation of new data being prioritized.

There is no lack of scientific and entrepreneurial talent in Africa, but there are fewer opportunities and fewer resources. Building up a critical mass in a discipline in Africa can be challenging, as our graduates and scientists are sought internationally and offered enticing positions with good salaries that cannot be matched locally. Although a testimony to scientific excellence and training quality in Africa, it presents challenges for growth and sustainability. This is exacerbated by a general paucity of academic and research opportunities in most African countries and, in some, a culture of stagnant hierarchy that may exclude young highfliers from climbing the academic ladder, acting together as deterrents for researchers to return to Africa and to develop their research careers in Africa.

Promoting equitable research partnerships could bring some balance. Success will be promoted by joint development of research questions and interpretation of results, by building in mechanisms to strengthen capacity to develop the next generation of scientists, and by careful consideration of resource and benefit sharing. International collaborations need to develop decision-making processes that give equal weight to partners and provide similar access to potential benefits. Datasets need to be jointly governed and opportunities for downstream analyses maximized with meaningful inputs from the African partners.

More attention is necessary for data harmonization to facilitate interoperability, and good data quality and reproducibility are essential components for data sharing. Privacy and risk of re-identification using personal data should be as-

sessed and measures introduced to minimize identifiability that could cause potential harm.

To scale the impact of genomic data on translational research in Africa will require capacity and infrastructure strengthening, dedicated support, political will, and public education and awareness.⁵ However, in the absence of more data and data sharing, the benefits will continue to elude the peoples of Africa, and we will have failed in our responsibility as researchers.

Sustainability of data sharing for a better future

There is a need to scale capacity for generating quality health data in African populations with careful thought to the different types of data that will inform clinically useful applications. Africa is lagging in electronic health records and large longitudinal cohorts with layered data to develop algorithms to predict the risk for developing non-communicable diseases. Extrapolations and predictions are often based on very small studies in specific target communities and have low generalizability yet are used to inform priority setting in many African regions. The diversity in climate, diet, socioeconomic environment, and culture necessitate research across African regions and in the context of the spectrum from rural to urban communities. The effects of genetic diversity, impacted by demographic history, natural selection and other evolutionary forces, impact health and need further investigation.⁶

A new NIH-funded research initiative, Harnessing Data Science for Health Discovery and Innovation in Africa (<https://commonfund.nih.gov/africadata/fundedresearch>), aims to promote health by using existing African data to explore analysis approaches and integration of diverse datasets; to strengthen capacity in machine learning and artificial intelligence; and to understand ethical, cultural and legal implications in different African countries and settings. Critical to this initiative is the foundational research from the Human Heredity and Health in Africa Consortium that has built African research networks and genetic data from over 70,000 African participants.

African genomes are an important key toward unlocking gene-environment interactions and interpreting the impact of genetic variation on the phenotype. We

need more data from Africans living in Africa, and we need enabling mechanisms to ensure that data sharing will also benefit those most affected by poverty, poor health, and low access to resources. Globalization of science and the potential for translational research through data sharing will benefit all people, irrespective of their origins or where they live. Doing this responsibly is the only way to ensure that no one is left behind.

ABOUT THE AUTHOR

Michèle Ramsay (PhD) is director of the Sydney Brenner Institute for Molecular Bioscience (SBIMB), professor in human genetics, and South African research chair in genomics and bioinformatics of African populations, based at the Univer-

sity of the Witwatersrand in Johannesburg. Her research aims to shed light on the role of genomic variation and environmental diversity in susceptibility to disease among African populations. While active in the leadership of the Human Heredity and Health in Africa Consortium and the International Hundred-thousand+ Cohorts Consortium, she advocates for responsible data sharing. Ramsay mentors young African scientists to promote capacity strengthening for genomic research and genetic services in African settings.

REFERENCES

1. Choudhury, A., Aron, S., Botigué, L.R., Sengupta, D., Botha, G., Bensellak, T., Wells, G., Kumuthini, J., Shriner, D., Fakim, Y.J., et al.; TrypanoGEN Research Group; H3Africa Consortium (2020). High-depth African genomes inform human migration and health. *Nature* 586, 741–748.
2. Fernando, B., King, M., and Sumathipala, A. (2019). Advancing good governance in data sharing and biobanking - international aspects. *Wellcome Open Res.* 4, 184.
3. Anane-Sarpong, E., Wangmo, T., and Tanner, M. (2020). Ethical principles for promoting health research data sharing with sub-Saharan Africa. *Developing World Bioeth.* 20, 86–95.
4. Munung, N.S., de Vries, J., and Pratt, B. (2021). Genomics governance: advancing justice, fairness and equity through the lens of the African communitarian ethic of Ubuntu. *Med. Health Care Philos.* 24, 377–388.
5. Kamp, M., Krause, A., and Ramsay, M. (2021). Has translational genomics come of age in Africa? *Hum. Mol. Genet.* 30, R164–R173.
6. Pereira, L., Mutesa, L., Tindana, P., and Ramsay, M. (2021). African genetic diversity and adaptation inform a precision medicine agenda. *Nat. Rev. Genet.* 22, 284–306.