frontiers | Research Topics

The ethics and challenges of studying the genetics of marginalized populations

Edited by
Arslan A. Zaidi, Keolu Fox, Katrina Claw and Tina Lasisi

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-3424-3
DOI 10.3389/978-2-8325-3424-3

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The ethics and challenges of studying the genetics of marginalized populations

Topic editors

Arslan A. Zaidi — University of Pennsylvania, United States
Keolu Fox — University of California, San Diego, United States
Katrina Claw — University of Colorado, United States
Tina Lasisi — University of Michigan, United States

Citation

Zaidi, A. A., Fox, K., Claw, K., Lasisi, T., eds. (2023). *The ethics and challenges of studying the genetics of marginalized populations*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3424-3

Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 22 August 2023
DOI 10.3389/fgene.2023.1260457

OPEN ACCESS

EDITED AND REVIEWED BY
Dov Greenbaum,
Yale University, United States

*CORRESPONDENCE
Tina Lasisi,
tlasisi@umich.edu
Arslan A. Zaidi,
szaidi@umn.edu

RECEIVED 17 July 2023
ACCEPTED 24 July 2023
PUBLISHED 22 August 2023

CITATION
Lasisi T, Claw KG, Fox K and Zaidi AA (2023), Editorial: The ethics and challenges of studying the genetics of marginalized populations.
Front. Genet. 14:1260457.
doi: 10.3389/fgene.2023.1260457

Editorial: The ethics and challenges of studying the genetics of marginalized populations

Tina Lasisi[1]*, Katrina G. Claw[2], Keolu Fox[3] and Arslan A. Zaidi[4]*

[1]Department of Anthropology, University of Michigan, Ann Arbor, MI, United States, [2]Division of Biomedical Informatics and Personalized Medicine, Colorado Center for Personalized Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States, [3]Department of Anthropology, Global Health Program, University of California, San Diego, La Jolla, CA, United States, [4]Genetics, Cell, and Developmental Biology Department, Institute of Health Informatics, University of Minnesota, Saint Paul, MN, United States

KEYWORDS

genetics, genomics, ethics, marginalized/vulnerable population, diversity and inclusion

Editorial on the Research Topic
The ethics and challenges of studying the genetics of marginalized populations

This Research Topic, entitled "*The Ethics and Challenges of Studying the Genetics of Marginalized Populations*," offers a pointed examination of the ethical considerations embedded within the study of genetics in historically underrepresented groups. Although there is much discourse on the need to include data from historically marginalized groups in genetics research, we believe that true inclusion lies not only in the diversification of samples but also that of researchers. As such, we have made a concerted effort in this Research Topic to invite contributors from diverse backgrounds.

Martschenko and Young open the discussion by challenging precision medicine's overreliance on broad racial and ethnic classifications. Their provocative critique reminds us that until our collective approach to precision medicine fully engages the complexity and full spectrum of human diversity, it will continue to fall short.

Silva et al.'s case study on the intersection of genomics, mestizaje, and Indigenous identities in Chile further deepens the discourse around the complexities of identity and representation in genetic research. Their cautionary narrative underscores the potential dangers of identity erasure and fetishization of indigeneity, urging the research community to maintain sensitivity towards the social complexities entangled in the field of genomics.

Taking a more environmental lens, Thompson and Crocker critique the common predilection for a genome-focused approach in health studies. They argue that significant health disparities are rooted in environmental factors rather than genetic differences, urging for a comprehensive approach that addresses these social and environmental aspects.

Through their practical work in Zamboanga and the Sulu Archipelago, Rodriguez et al. elucidate the unique ethical challenges associated with conducting genetic research among Indigenous Peoples. Their important work underscores the necessity of respecting the sovereignty of Indigenous Peoples over their genetic information and of developing respectful, equitable research partnerships.

Villanea and Witt address the challenging Research Topic of underrepresented populations in archaic introgression research. Their insightful piece raises vital concerns about potential inequities in research design, highlighting the necessity for equitable research practices and fair benefit sharing.

Jackson et al. provide a necessary critique regarding the underrepresentation of African Americans in genomic studies and present compelling case studies that illustrate the pressing necessity of ensuring broad and fair representation in genetic research.

In their exploration of data sovereignty, Carroll et al. propose the crucial need for Indigenous standards in control and oversight of biomedical data, offering a distinct perspective on the implementation of the CARE Principles for Indigenous Data Governance.

Finally, Young et al. view the challenge of recruiting diverse participants in genomics research through a nuanced lens, highlighting the influence of sociodemographic factors on research participation and emphasizing the importance of various recruitment strategies to ensure diversity.

This set of insightful articles offers a rich palette of perspectives on the myriad ethical challenges faced in genetic studies involving historically marginalized populations. Each piece, while unique in its approach, unites in its mission to shed light on these challenges and deliver potential solutions.

These contributions have critical implications for the wider scientific community, both in terms of advancing our understanding of the genetic architecture of disease risk specifically for non-European (descent) populations, and in highlighting the ethical challenges and proposed solutions for involving and studying marginalized populations. This collective body of work enriches our scientific dialogue, urging us towards more ethical and inclusive research practices.

The contributions illuminate a multifaceted picture of the current status of diverse genetic studies and advocate for a deeper engagement with marginalized populations. The sobering reality of ongoing inequities in the field underscores reliance on past models of research, which are unsatisfactory, and urges a shared vision for the future, where genetic studies are representative, ethical, and equitable.

Importantly, these discussions emphasize the importance of meaningful community involvement, transparent communication, and respect for cultural histories and identities. As we move forward, we must ensure that all populations can participate in, contribute to, and potentially benefit from advances in genetic research. This work underscores the compelling need to continually revisit and reconsider the ethics and practices not only in genetics but in all scientific fields engaging with marginalized populations.

This Research Topic can serve as a foundation to broaden and deepen future research efforts, ultimately enhancing and enlightening our collective understanding of human health and disease.

Author contributions

TL: Writing-original draft, Writing-review and editing. KC: Writing-review and editing. KF: Writing-review and editing. AZ: Writing-review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Genetics

OPINION
published: 25 February 2022
doi: 10.3389/fgene.2022.821170

Underrepresented Populations at the Archaic Introgression Frontier

Fernando A Villanea[1]* *and Kelsey E. Witt*[2]

[1]Department of Anthropology, College of Arts and Sciences, University of Colorado Boulder, Boulder, CO, United States, [2]Department of Ecology and Evolutionary Biology, Brown University, Providence, RI, United States

Keywords: ethics, underrepresented populations, archaic admixture, anthropology, human genomics

INTRODUCTION

Recent advancements in the recovery of ancient genomes have yielded high-coverage sequences for two archaic human species: Neanderthals and Denisovans. Perhaps more surprisingly, direct comparisons of archaic and modern human genomes have revealed a complex landscape of admixture between both archaic species and modern humans (Browning et al., 2018; Villanea and Schraiber, 2019). While we call these regions of the human genome "archaically introgressed", they are functional contributors to the living human gene pool, affecting our health and fitness. For Neanderthals in particular, early archaic ancestry maps focused on modern Eurasians, as hundreds of genomes from Europe and East Asia were readily available from the 1000 Genomes Project (see Sankararaman et al., 2014; Vernot and Akey, 2014). Coupled with the geographic distribution of Neanderthal archaeological sites, which are largely located in Europe, this created a strong impression to the larger public that individuals of European descent, in particular, carried archaic genomic elements, which coincided with a larger interest in commodifying archaic ancestry by personalized genomic companies—as evidenced by 23 and Me incorporating a report of Neanderthal ancestry into their mainline product.

As scientific efforts progressed to identify regions of the modern human genome that originated in these archaic populations, in a twist of irony, European populations were found to retain the smallest component of both Neanderthal and Denisovan genome ancestry outside of African populations (Sankararaman et al., 2016). Current archaic genome studies indicate that South Asian populations, as defined by the 1000 Genomes Project (which encompasses populations from Bangladesh, India, Pakistan, and Sri Lanka), have a larger component of Neanderthal ancestry than Europeans, East Asians, or Maritime Southeast Asians (Witt et al., 2021 bioRXiv, **Figure 1**). Conversely, Maritime Southeast Asian populations retain a larger component of the genome, and more unique variation, of Denisovan ancestry (Sankararaman et al., 2016; Vernot et al., 2016, **Figure 1**). As studies continued to identify regions of the human genome enriched for archaic ancestry at the population level, a consistent pattern emerged for the distribution of archaic ancestry: for a vast majority of genes found in modern humans, archaic variants appear to have been removed from the gene pool by negative natural selection (Harris and Nielson, 2016; Petr et al., 2019), but there is also enrichment in a minority of functional regions that indicate positive selection (Racimo et al., 2017; Jagoda et al., 2018). Thus, the focus has shifted from viewing archaic ancestry as a quirk of human evolution into understanding the functional importance of these rare genomic regions enriched for archaic ancestry.

FUNCTIONAL IMPORTANCE OF ARCHAIC ANCESTRY

There are multiple examples of archaic gene variants that helped modern humans to adapt to novel environments as they expanded throughout the world, such as differences in UV radiation exposure,

OPEN ACCESS

Edited by:
Tina Lasisi,
The Pennsylvania State University (PSU), United States

Reviewed by:
Omer Gokcumen,
University at Buffalo, United States
Maria A. Nieves Colón,
University of Minnesota Twin Cities, United States
Aylwyn Scally,
University of Cambridge, United Kingdom

***Correspondence:**
Fernando A Villanea
fernando.villanea@colorado.edu

Specialty section:
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

Received: *23 November 2021*
Accepted: *07 February 2022*
Published: *25 February 2022*

Citation:
Villanea FA and Witt KE (2022) Underrepresented Populations at the Archaic Introgression Frontier. Front. Genet. 13:821170. doi: 10.3389/fgene.2022.821170

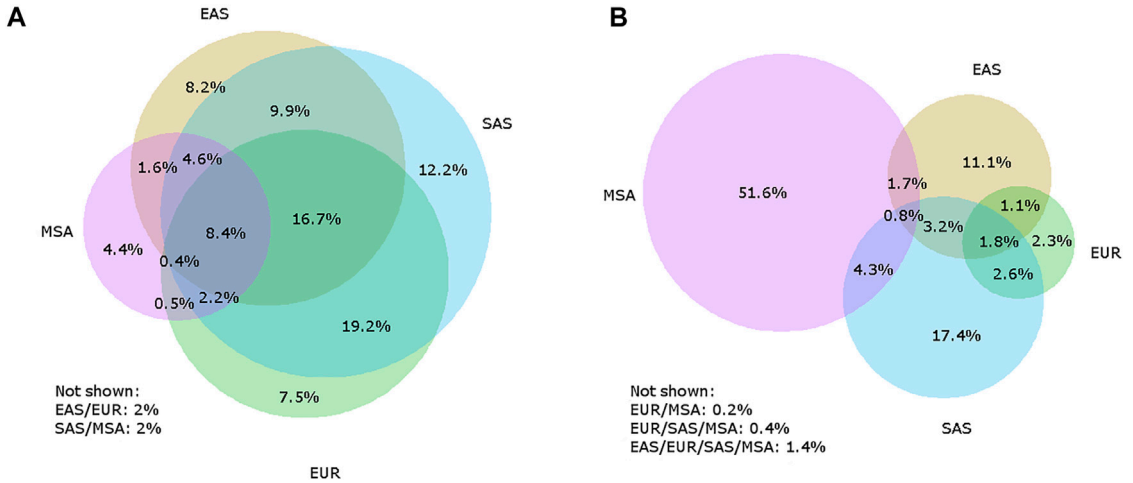


FIGURE 1 | Venn diagrams illustrating how archaic variants are shared between Europeans (EUR), East Asians (EAS), and South Asians (SAS) as defined by the 1000 Genomes Project, and Maritime Southeast Asians (MSA) from the Simons Genome Diversity Project for Neanderthal-unique **(A)** and Denisovan-unique **(B)** variants. The listed percentages are the percent of the total number of Neanderthal-unique or Denisovan-unique variants. It is not always possible to accurately display all possible overlaps with a four-circle Venn diagram, so any sets that are not shown have percentages included below each Venn diagram, with the circles overlapping listed. For example, "EAS/EUR" refers to all variants that are shared between East Asians and Europeans but are not found in South Asians or Maritime Southeast Asians.

temperature, dietary composition, and pathogen exposure. Functional genes enriched for archaic ancestry include the Neanderthal BNC2 and OCA2 variants related to skin pigmentation (Sankararaman et al., 2014; Vernot and Akey 2015; Gittelman et al., 2016); the OAS locus and Toll-like receptor loci related to immune response (Mendez et al., 2012; Dannemann and Kelso 2016; Gittelman et al., 2016; Sams et al., 2016); and the TBX15/WARS2 locus related to lipid metabolism (Racimo et al., 2017). Furthermore, there is mounting evidence that the bulk of contributions to modern human fitness and health from archaic ancestry is through non-gene functional portions of the human genome, which are far more difficult to conceptualize (Dannemann et al., 2017; Petr et al., 2019; Silvert et al., 2019). For example, some enhancer regions show enrichment in Neanderthal alleles, such as adipose-related tissues and primary T-cells (Dannemann et al., 2017; Petr et al., 2019; Silvert et al., 2019), plus, up to an additional forty-two tissues in humans show significant enrichment of archaic alleles in enhancers, with the highest rate of enrichment identified in adipose-related tissues and immune cells (Silvert et al., 2019).

Adaptive introgression seems particularly prominent in immune-related functional genome elements, suggesting that Neanderthals and Denisovans harbored many alleles adapted to local pathogens that were positively selected after introgression into modern humans (Ahlquist et al., 2021). Genes that encode proteins that interact with RNA viruses are also enriched for introgressed alleles (Enard and Petrov, 2018). Similarly, polygenic adaptive introgression has been reported in pathways associated with immunity (Gouy and Excoffier 2020). Finally, population transcriptome studies of immune response to viral and bacterial pathogens have found many gene expression and splicing differences between individuals of European and African ancestry that appear to be driven by Neanderthal introgressed alleles, providing further support for their regulatory impact on immunity (Nedelec et al., 2016; Quach et al., 2016; Rotival et al., 2019). Recently, a study reporting the expression of Neanderthal non-gene elements found as many as 292 expression-modulating variants in human lineages, most of them related to immune function, underlining the importance of archaic variation in modulating the expression of modern human genes (Jagoda et al., 2021). Although the archaic populations that admixed with modern humans are now extinct, the archaic variation remaining in the human gene pool can have significant impacts on health, especially immune function, and therefore is an important target for genomics and biomedical research.

IMPORTANCE OF NON-EUROPEAN GENOMES FOR UNDERSTANDING ARCHAIC GENETIC VARIATION

The field of genomics has a well-known Eurocentric bias, where European populations are more broadly sampled and variation found in European populations is far better characterized than other populations (e.g. Popejoy and Fullerton, 2016). Interestingly, despite this thorough sampling, a larger potential for discovering novel adaptive archaic introgression exists in non-European genomes, as Europeans have some of the lowest proportions of both Neanderthal and Denisovan ancestry outside of African populations (Sankararaman et al., 2014, 2016).

Given the recent interest in identifying archaic functional variants in modern humans, a more effective study design is to examine diverse populations that have thrived in varied environmental conditions, to identify archaic variants that

may have been adaptive. For example, the most compelling case of adaptive introgression of an archaic gene to date is the high-altitude adaptation of Tibetans achieved through selection for the archaic variant of the EPAS1 gene, which was introduced into the ancestral Tibetan populations through admixture with Denisovans (Huerta-Sánchez et al., 2014; Zhang et al., 2021). This example highlights how genetic variation unique to Tibetans was paramount in understanding the role of archaic variation in modern humans in the first place, variation that simply does not exist in any other population. Another example of an adaptive archaic variant found in underrepresented populations is a Denisovan variant of TBX15, which is found at high frequency in Greenlandic Inuit (Racimo et al., 2018). This Denisovan haplotype is found in many Indigenous populations of the Americas at a high frequency (>0.8), and in Inuit specifically has a strong signal of positive selection. TBX15 is linked to a number of phenotypes, including lipid metabolism, especially at cold temperatures, which suggests that the Denisovan variant may have been adaptive for humans as they populated the Arctic. The archaic alleles are also in linkage disequilibrium with a SNP that is linked to waist-hip ratio, with an effect size of 0.034 (Heid et al., 2010; Fumagalli et al., 2015). This SNP was identified in European populations, but the haplotype in Greenlandic Inuit is not well-characterizing, underlining the need for further study of these underrepresented populations. These two examples suggest that the further study of populations that are generally underrepresented in genomic research, including Melanesians, Southeast Asians, and Indigenous American populations, could yield additional novel archaic functional variants.

While many of the early genetic analyses of archaic ancestry in modern humans have focused on continental Eurasian populations (especially Europeans and East Asians), Papuans and other populations in Oceania have a unique distribution of archaic ancestry (**Figure 1**). Papuans have a high proportion of Neanderthal ancestry, but most notably they have an extremely high proportion of Denisovan ancestry compared to other populations worldwide - in some populations, as high as 5% (Sankararaman et al., 2016; Vernot et al., 2016). Additionally, evidence suggests that Papuans have Denisovan ancestry from two genetically distinct Denisovan populations (Browning et al., 2018; Jacobs et al., 2019). This is evident as the majority of Denisovan variants in Papuans are not shared with Eurasian populations (**Figure 1**). Recently, the Ayta Magbukon people in the Philippines have been reported to possess Denisovan ancestry proportions even higher than that of Papuans (Larena et al., 2021). Southeast Asia also has a long history of occupation by multiple archaic hominins, including *Homo erectus* (Rizal et al., 2020) and *Homo floresiensis* (Brown et al., 2004; Sutikna et al., 2016), and possibly a new hominin, currently named *Homo luzonensis* (Detroit et al., 2019). Some of these hominins may have lived in the region around the same time period, and some likely overlapped with humans although evidence for gene flow between these hominins and modern humans has yet to be identified (Teixeira et al., 2021).

Another underrepresented group of populations with an interesting legacy of archaic ancestry are Indigenous Americans. Indigenous American populations today are the descendants of individuals who populated the American continent in a process that started at least 25,000 years before present (Bennet et al., 2021). Indigenous American individuals carry archaic genomic elements at frequencies comparable to modern East Asian individuals (Sankararaman et al., 2016), indicating that the first American migrants already carried archaic genomic elements. In the 25,000 or more years since these founding populations migrated to the Americas, these peoples would have encountered numerous novel environments, and adapting to meet those environmental challenges helped to shape their genomes into the unique populations living today. This long adaptation process would have undoubtedly also impacted the distribution of archaic genomic elements in Indigenous American populations.

The European colonization of much of the American continent has also left a profound impact on its inhabitants, visible in the genomes of all individuals today, as most modern Indigenous American individuals have both European ancestry as a result of colonization and African ancestry as a result of African individuals being forcibly relocated through the slave trade. This admixture has also affected the amount of archaic ancestry in modern Indigenous American populations: admixture with Europeans and Africans, which have slightly less and significantly less archaic ancestry respectively, resulted in a dilution of archaic ancestry in modern Indigenous American descendants (**Figure 2**). Natural selection has been an important force in shaping Indigenous American genomes, both pre-admixture (Williams et al., 2014; Reynolds et al., 2016) as well as post-admixture (Ongaro et al., 2021). Recent admixture between populations can have both positive and negative health consequences: for example, individuals with recent ancestors from multiple worldwide populations show reduced risk for some genetic disorders but an increased incidence of autoimmune disease (Rudan, 2006; Martin et al., 2017). Therefore, the dilution of archaic elements in modern Indigenous populations could provide health benefits, but could also have a negative impact on health: by replacement with maladaptive gene variants, by breaking up existing epistatic gene interactions, or by interactions between gene expression-modulating variants. Because of these changes to genetic architecture as a result of admixture following European contact, advances in personalized medicine for Indigenous American individuals could be extremely beneficial, yet any future endeavor requires addressing a long history of bad faith interactions between geneticists and Indigenous American communities.

While it is clear that the study of archaic introgression in modern humans would benefit from the analysis of populations that are often underrepresented in genomic research, a balance must be struck between the scientific desire for knowledge and respect for these populations and their right to control their own data. Past work by geneticists with marginalized communities has often been exploitative, involved little interaction between researchers and communities after samples were collected, and rarely resulted in benefits for the individuals being studied (e.g., Hart and Sobraske, 2003). This history of poor interactions between the scientific community and marginalized

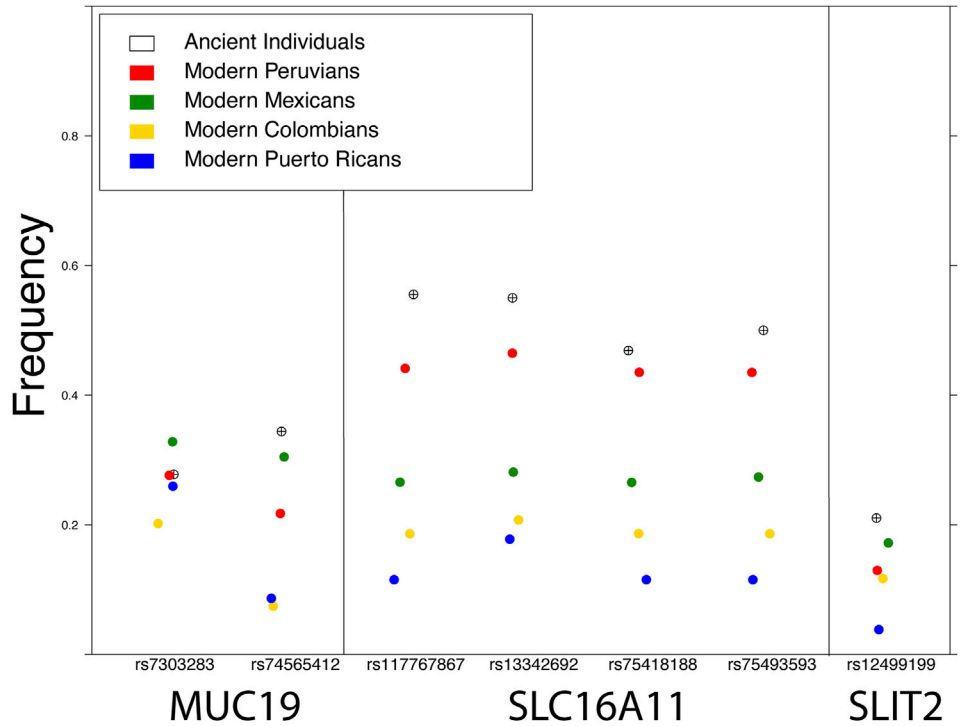


FIGURE 2 | Frequency of the archaic variant of SNPs in three genes reported under selection in modern Indigenous American populations: MUC19 (Racimo et al., 2016; Reynolds et al., 2019), SLC16A11 (Williams et al., 2014), and SLIT2 (Reynolds et al., 2019). The frequency in ancient individuals was calculated from various pre-European colonization genomes available in the literature (see Figure *Methods*).

populations has resulted in the erosion of trust in scientists, and some communities have rightfully placed a moratorium on genetic research to limit the infliction of future harms (Begay et al., 2020; Claw et al., 2021). Many Indigenous scholars have provided suggestions and guidelines to ensure that future genomic research is conducted in an equitable way that ensures communities are able to make their own decisions concerning their research and data, and also equipping them to conduct the research themselves (Claw et al., 2018; Fox, 2020; Hudson et al., 2020; Tsosie et al., 2020).

Community organization is key for maintaining data sovereignty and creating leverage for the distribution of material benefits arising from unique genomic variation. The Khoisan communities of South Africa have published the San Code of Research Ethics (globalcodeofconduct.org/affiliated-codes/), in which they collectively dictate how they will interact with researchers. Global measures implemented to help ensure communities being studied benefit from the material returns of genetic research include the "Nagoya Protocol on Access and Benefit-sharing" (cbd.int/abs/); an international agreement that established guidelines for "sharing the benefits arising from the utilization of genetic resources in a fair and equitable way", and as of November 2020 the agreement has been ratified or accessioned by 132 countries. Other communities have also chosen to control their own genetic resources. For example, the Native BioData Consortium (NBDC, https://nativebio.org/), was founded in 2018 by Indigenous scientists as a biobank for Indigenous communities. The NBDC also conducts its own research projects and hosts training workshops for Indigenous students. Further examples of community interests built into genomic projects include initiatives such as Variant Bio (variantbio.com/), a company which forms partnerships with communities possessing unique genetic diversity, builds the priorities of the communities into their study designs, and is committed to redistributing royalties from any medically important discoveries originating from their genomes.

In the near future, research into further evolutionary adaptations born of archaic genome elements will drive the need to genotype other underrepresented populations globally. As the funding available for these studies is largely Eurocentric, but the genetic diversity of interest exists outside of Europeans, this may result in exploitative research designs. Because of this asymmetry, it is paramount to build scientific capacity within these communities, and prioritize funding for existing scientists of underrepresented populations. At the same time, community engagement needs to be built into all levels of study designs; research and funding schemes must be intrinsically flexible to accommodate the priorities of the communities co-designing the studies.

Likewise, ultimate control of data sharing should be managed directly by community-lead organizations. While open science has been championed as an ideal model, the current application of open science has neither benefited Indigenous communities,

nor is it particularly "open" as it is currently implemented. As an example, publicly available human genomic data is already compartmentalized between repositories such as the 1000 Genomes Project, the Simons Genome Diversity Project, and the Human Genome Diversity Project, among other more specialized databases, in addition to research groups that host and maintain their own data, such as the Neanderthal and Denisovan genomes from the Max Planck Institute for Evolutionary Anthropology. Each one of these entities determines what permissions are required to access their data, meaning that in practice, there is not one centralized way to access human genomic data. Going forward, the addition of community-led data management would maintain similar compartmentalization as exists today, but provide Indigenous communities with leverage in negotiations on how that data is used (Mc Cartney et al., 2022). As legal protections and community engagement procedures are more well-established in some regions of the world than others, the community-involved research should be conducted not just with respect to local laws, but in a way that truly puts decision-making in the hands of the populations being studied. Finally, the format for the dissemination of results should recognize the communication traditions of participating communities to maximize digestibility of scientific and medical findings, to ensure transparency and informed consent during every study step. This transparency in reporting should also explicitly inform the financial potential for practical applications of new findings, to provide leverage for the negotiations of royalties from any medically important discoveries stemming from their unique genomic ancestry.

FIGURE METHODS

Figure 1 To identify sites containing archaic alleles, we used a method previously employed by Witt et al. (2021, biorXiv) that considers archaic alleles to be shared with modern humans and Neanderthals/Denisovans, present in African populations at a frequency of less than 1%, and present in a population outside of Africa with a frequency of at least 1%. We used the 1000 Genomes Project (1KGP) genomic data for our African comparison population and identified archaic alleles in the non-African 1KGP populations as well as Papuans, whose genomes were collected as part of the Simons Genome Diversity Project (SGDP). We classified those populations into regions (Americas, Europe, East Asia, South Asia, and Papua New Guinea, as defined by the 1KGP and SGDP, and determined which positions in the genome contained archaic alleles for each region. To examine the patterns of variant sharing between Neanderthals and Denisovans separately, we only considered variants present in one population and absent from the other, which we term Neanderthal-unique and Denisovan-unique variants.

Figure 2 Modern frequencies were calculated from four populations publicly available in the 1000 Genomes Project. Frequencies in ancient individuals were calculated by combining high coverage (>1X) pre-European contact genomes from the literature, including nine individuals from California and one from Ontario (Scheib et al., 2018), four from Peru (Lindo et al., 2018), four from Patagonia (de la Fuente et al., 2018), one from Alaska (Moreno-Mayar et al., 2018), and one from Montana (Rasmussen et al., 2014).

AUTHOR CONTRIBUTIONS

Both authors conceived and wrote this manuscript. KW generated the data for **Figure 1**, FV generated the data for **Figure 2**.

ACKNOWLEDGMENTS

The authors would like to thank David Peede for making the data for **Figure 2** available for this manuscript.

REFERENCES

Ahlquist, K., Bañuelos, M. M., Funk, A., Lai, J., Rong, S., Villanea, F. A., et al. (2021). Our Tangled Family Tree: New Genomic Methods Offer Insight into the Legacy of Archaic Admixture. *Genome Biol. Evol.* 13, evab115. doi:10.1093/gbe/evab115

Begay, R. L., Garrison, N. A., Sage, F., Bauer, M., Knoki-Wilson, U., Begay, D. H., et al. (2020). Weaving the Strands of Life (Iiná Bitł'ool): History of Genetic Research Involving Navajo People. *Hum. Biol.* 91 (3), 189–208. doi:10.13110/humanbiology.91.3.04

Bennett, M. R., Bustos, D., Pigati, J. S., Springer, K. B., Urban, T. M., Holliday, V. T., et al. (2021). Evidence of Humans in North America during the Last Glacial Maximum. *Science* 373 (6562), 1528–1531. doi:10.1126/science.abg7586

Brown, P., Sutikna, T., Morwood, M. J., Soejono, R. P., Saptomo, E. W., and Due, R. A. (2004). A New Small-Bodied Hominin from the Late Pleistocene of Flores, Indonesia. *Nature* 431 (7012), 1055–1061. doi:10.1038/nature02999

Browning, S. R., Browning, B. L., Zhou, Y., Tucci, S., and Akey, J. M. (2018). Analysis of Human Sequence Data Reveals Two Pulses of Archaic Denisovan Admixture. *Cell* 173 (1), 53–61. doi:10.1016/j.cell.2018.02.031

Claw, K. G., Anderson, M. Z., Begay, R. L., Tsosie, K. S., Fox, K., and Garrison, N. A. (2018). A Framework for Enhancing Ethical Genomic Research with Indigenous Communities. *Nat. Commun.* 9 (1), 2957–7. doi:10.1038/s41467-018-05188-3

Claw, K. G., Dundas, N., Parrish, M. S., Begay, R. L., Teller, T. L., Garrison, N. A., et al. (2021). Perspectives on Genetic Research: Results from a Survey of Navajo Community Members. *Front. Genet.* 12, 1–11. doi:10.3389/fgene.2021.734529

Dannemann, M., and Kelso, J. (2017). The Contribution of Neanderthals to Phenotypic Variation in Modern Humans. *Am. J. Hum. Genet.* 101, 578–589. doi:10.1016/j.ajhg.2017.09.010

de la Fuente, C., Ávila-Arcos, M. C., Galimany, J., Carpenter, M. L., Homburger, J. R., Blanco, A., et al. (2018). Genomic Insights into the Origin and Diversification of Late Maritime hunter-gatherers from the Chilean Patagonia. *Proc. Natl. Acad. Sci. USA* 115 (17), E4006–E4012. doi:10.1073/pnas.1715688115

Détroit, F., Mijares, A. S., Corny, J., Daver, G., Zanolli, C., Dizon, E., et al. (2019). A New Species of Homo from the Late Pleistocene of the Philippines. *Nature* 568, 181–186. doi:10.1038/s41586-019-1067-9

Enard, D., and Petrov, D. A. (2018). Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. *Cell* 175, 360–371. doi:10.1016/j.cell.2018.08.034

Fox, K. (2020). The Illusion of Inclusion - the "All of Us" Research Program and Indigenous Peoples' DNA. *N. Engl. J. Med.* 383 (5), 411–413. doi:10.1056/nejmp1915987

Fumagalli, M., Moltke, I., Grarup, N., Racimo, F., Bjerregaard, P., Jørgensen, M. E., et al. (2015). Greenlandic Inuit Show Genetic Signatures of Diet and Climate Adaptation. *Science* 349 (6254), 1343–1347. doi:10.1126/science.aab2319

Gittelman, R. M., Schraiber, J. G., Vernot, B., Mikacenic, C., Wurfel, M. M., and Akey, J. M. (2016). Archaic Hominin Admixture Facilitated Adaptation to Out-Of-Africa Environments. *Curr. Biol.* 26, 3375–3382. doi:10.1016/j.cub.2016.10.041

Gouy, A., and Excoffier, L. (2020). Polygenic Patterns of Adaptive Introgression in Modern Humans Are Mainly Shaped by Response to Pathogens. *Mol. Biol. Evol.* 37, 1420–1433. doi:10.1093/molbev/msz306

Harris, K., and Nielsen, R. (2016). The Genetic Cost of Neanderthal Introgression. *Genetics* 203 (2), 881–891. doi:10.1534/genetics.116.186890

Hart, S., and Keith, S. (2003). *Investigative Report Concerning the Medical Genetics Project at Havasupai*. Phoenix, AZ, USA: Arizona State University Law Library.

Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Steinthorsdottir, V., et al. (2010). Meta-analysis Identifies 13 New Loci Associated with Waist-Hip Ratio and Reveals Sexual Dimorphism in the Genetic Basis of Fat Distribution. *Nat. Genet.* 42 (11), 949–960. doi:10.1038/ng.685

Hudson, M., Garrison, N. A., Sterling, R., Caron, N. R., Fox, K., Yracheta, J., et al. (2020). Rights, Interests and Expectations: Indigenous Perspectives on Unrestricted Access to Genomic Data. *Nat. Rev. Genet.* 21 (6), 377–384. doi:10.1038/s41576-020-0228-x

Huerta-Sánchez, E., Jin, X., Bianba, Z., Peter, B. M., Vinckenbosch, N., Liang, Y., et al. (2014). Altitude Adaptation in Tibetans Caused by Introgression of Denisovan-like DNA. *Nature* 512 (7513), 194–197.

Jacobs, G. S., Hudjashov, G., Saag, L., Kusuma, P., Darusallam, C. C., Lawson, D. J., et al. (2019). Multiple Deeply Divergent Denisovan Ancestries in Papuans. *Cell* 177 (4), 1010–1021. doi:10.1016/j.cell.2019.02.035

Jagoda, E., Xue, J. R., Reilly, S. K., Dannemann, M., Racimo, F., Huerta-Sanchez, E., et al. (2021). Detection of Neanderthal Adaptively Introgressed Genetic Variants that Modulate Reporter Gene Expression in Human Immune Cells. *Mol. Biol. Evol.* doi:10.1093/molbev/msab304

Jagoda, E., Lawson, D. J., Wall, J. D., Lambert, D., Muller, C., Westaway, M., et al. (2018). Disentangling Immediate Adaptive Introgression from Selection on Standing Introgressed Variation in Humans. *Mol. Biol. Evol.* 35 (3), 623–630. doi:10.1093/molbev/msx314

Larena, M., McKenna, J., Sanchez-Quinto, F., Bernhardsson, C., Ebeo, C., Reyes, R., et al. (2021). Philippine Ayta Possess the Highest Level of Denisovan Ancestry in the World. *Curr. Biol.* 31 (19), 4219–4230. doi:10.1016/j.cub.2021.07.022

Lindo, J., Haas, R., Hofman, C., Apata, M., Moraga, M., Verdugo, R. A., et al. (2018). The Genetic Prehistory of the Andean highlands 7000 Years BP Though European Contact. *Sci. Adv.* 4 (11), eaau4921. doi:10.1126/sciadv.aau4921

Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., et al. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* 100, 635–649. doi:10.1016/j.ajhg.2017.03.004

Mc Cartney, A. M., Anderson, J., Liggins, L., Hudson, M. L., Anderson, M. Z., TeAika, B., et al. (2022). Balancing Openness with Indigenous Data Sovereignty: An Opportunity to Leave No One behind in the Journey to Sequence All of Life. *Proc. Natl. Acad. Sci.* 119, 4. doi:10.1073/pnas.2115860119

Mendez, F. L., Watkins, J. C., and Hammer, M. F. (2012). A Haplotype at STAT2 Introgressed from Neanderthals and Serves as a Candidate of Positive Selection in Papua New Guinea. *Am. J. Hum. Genet.* 91, 265–274. doi:10.1016/j.ajhg.2012.06.015

Moreno-Mayar, J. V., Potter, B. A., Vinner, L., Steinrücken, M., Rasmussen, S., Terhorst, J., et al. (2018). Terminal Pleistocene Alaskan Genome Reveals First Founding Population of Native Americans. *Nature* 553 (7687), 203–207. doi:10.1038/nature25173

Nédélec, Y., Sanz, J., Baharian, G., Szpiech, Z. A., Pacis, A., Dumaine, A., et al. (2016). Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. *Cell* 167, 657–669. e21.

Ongaro, L., Mondal, M., Flores, R., Marnetto, D., Molinaro, L., Alarcón-Riquelme, M. E., et al. (2021). Continental-scale Genomic Analysis Suggests Shared post-admixture Adaptation in the Americas. *Hum. Mol. Genet.* 30 (22), 2123–2134. doi:10.1093/hmg/ddab177

Petr, M., Pääbo, S., Kelso, J., and Vernot, B. (2019). Limits of Long-Term Selection against Neandertal Introgression. *Proc. Natl. Acad. Sci. USA* 116 (5), 1639–1644. doi:10.1073/pnas.1814338116

Popejoy, A. B., and Fullerton, S. M. (2016). Genomics Is Failing on Diversity. *Nature* 538, 161–164. doi:10.1038/538161a

Quach, H., Rotival, M., Pothlichet, J., Loh, Y.-H. E., Dannemann, M., Zidane, N., et al. (2016). Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell* 167, 643–656. e17. doi:10.1016/j.cell.2016.09.024

Racimo, F., Gokhman, D., Fumagalli, M., Ko, A., Hansen, T., Moltke, I., et al. (2018). Archaic Adaptive Introgression in TBX15/WARS2. *Mol. Biol. Evol.* 34 (3), 509–524. doi:10.1093/molbev/msw283

Racimo, F., Marnetto, D., and Huerta-Sánchez, E. (2017). Signatures of Archaic Adaptive Introgression in Present-Day Human Populations. *Mol. Biol. Evol.* 34 (2), 296–317. doi:10.1093/molbev/msw216

Rasmussen, M., Anzick, S. L., Waters, M. R., Skoglund, P., DeGiorgio, M., Stafford, T. W., et al. (2014). The Genome of a Late Pleistocene Human from a Clovis Burial Site in Western Montana. *Nature* 506 (7487), 225–229. doi:10.1038/nature13025

Reynolds, A. W., Mata-Míguez, J., Miró-Herrans, A., Briggs-Cloud, M., Sylestine, A., Barajas-Olmos, F., et al. (2019). Comparing Signals of Natural Selection between Three Indigenous North American Populations. *Proc. Natl. Acad. Sci. USA* 116 (19), 9312–9317. doi:10.1073/pnas.1819467116

Rizal, Y., Westaway, K. E., Zaim, Y., van den Bergh, G. D., Bettis, E. A., Morwood, M. J., et al. (2020). Last Appearance of Homo Erectus at Ngandong, Java, 117,000-108,000 Years Ago. *Nature* 577, 381–385. doi:10.1038/s41586-019-1863-2

Rotival, M., Quach, H., and Quintana-Murci, L. (2019). Defining the Genetic and Evolutionary Architecture of Alternative Splicing in Response to Infection. *Nat. Commun.* 10, 1971. doi:10.1038/s41467-019-09689-7

Rudan, I. (2006). Health Effects of Human Population Isolation and Admixture. *Croat. Med. J.* 47 (4), 526–531.

Sams, A. J., Dumaine, A., Nédélec, Y., Yotova, V., Alfieri, C., Tanner, J. E., et al. (2016). Adaptively Introgressed Neandertal Haplotype at the OAS Locus Functionally Impacts Innate Immune Responses in Humans. *Genome Biol.* 17 (1), 246. doi:10.1186/s13059-016-1098-6

Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S., et al. (2014). The Genomic Landscape of Neanderthal Ancestry in Present-Day Humans. *Nature* 507 (7492), 354–357. doi:10.1038/nature12961

Sankararaman, S., Mallick, S., Patterson, N., and Reich, D. (2016). The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Curr. Biol.* 26 (9), 1241–1247. doi:10.1016/j.cub.2016.03.037

Scheib, C. L., Li, H., Desai, T., Link, V., Kendall, C., Dewar, G., et al. (2018). Ancient Human Parallel Lineages within North America Contributed to a Coastal Expansion. *Science* 360 (6392), 1024–1027. doi:10.1126/science.aar6851

Silvert, M., Quintana-Murci, L., and Rotival, M. (2019). Impact and Evolutionary Determinants of Neanderthal Introgression on Transcriptional and Post-Transcriptional Regulation. *Am. J. Hum. Genet.* 104, 1241–1250. doi:10.1016/j.ajhg.2019.04.016

Sutikna, T., Tocheri, M. W., Morwood, M. J., Saptomo, E. W., Awe, R. D., Wasisto, S., et al. (2016). Revised Stratigraphy and Chronology for Homo Floresiensis at Liang Bua in Indonesia. *Nature* 532 (7599), 366–369. doi:10.1038/nature17179

Teixeira, J. C., Jacobs, G. S., Stringer, C., Tuke, J., Hudjashov, G., Purnomo, G. A., et al. (2021). Widespread Denisovan Ancestry in Island Southeast Asia but No Evidence of Substantial Super-archaic Hominin Admixture. *Nat. Ecol. Evol.* 5, 616–624. doi:10.1038/s41559-021-01408-0

Tsosie, K. S., Begay, R. L., Fox, K., and Garrison, N. A. (2020). Generations of Genomes: Advances in Paleogenomics Technology and Engagement for Indigenous People of the Americas. *Curr. Opin. Genet. Develop.* 62, 91–96. doi:10.1016/j.gde.2020.06.010

Vernot, B., and Akey, J. M. (2014). Resurrecting Surviving Neandertal Lineages from Modern Human Genomes. *Science* 343 (6174), 1017–1021. doi:10.1126/science.1245938

Vernot, B., Tucci, S., Kelso, J., Schraiber, J. G., Wolf, A. B., Gittelman, R. M., et al. (2016). Excavating Neandertal and Denisovan DNA from the Genomes of Melanesian Individuals. *Science* 352 (6282), 235–239. doi:10.1126/science.aad9416

Vernot, B., and Akey, J. M. (2015). Complex History of Admixture between Modern Humans and Neandertals. *Am. J. Hum. Genet.* 96, 448–453. doi:10.1016/j.ajhg.2015.01.006

Villanea, F. A., and Schraiber, J. G. (2019). Multiple Episodes of Interbreeding between Neanderthal and Modern Humans. *Nat. Ecol. Evol.* 3 (1), 39–44. doi:10.1038/s41559-018-0735-8

Williams, A. L., Williams, A. L., Jacobs, S. B., Moreno-Macías, H., Huerta-Chagoya, A., Churchhouse, C., et al. (2014). Sequence Variants in SLC16A11 Are a

Common Risk Factor for Type 2 Diabetes in Mexico. *Nature* 506 (7486), 97–101. doi:10.1038/nature12828

Witt, K. E., Villanea, F., Loughran, E., and Huerta-Sanchez, E. (2021). On the Apportionment of Archaic Human Diversity. *bioRxiv*.

Zhang, X., Witt, K. E., Bañuelos, M. M., Ko, A., Yuan, K., Xu, S., et al. (2021). The History and Evolution of the Denisovan-EPAS1 Haplotype in Tibetans. *Proc. Natl. Acad. Sci.* 118, 22. doi:10.1073/pnas.2020803118

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

OPINION
published: 03 March 2022
doi: 10.3389/fgene.2022.817318

The Articulation of Genomics, Mestizaje, and Indigenous Identities in Chile: A Case Study of the Social Implications of Genomic Research in Light of Current Research Practices

Constanza P. Silva[1,2], Constanza de la Fuente Castro[3]*, Tomás González Zarzar[4,5]*, Maanasa Raghavan[3], Ayelén Tonko-Huenucoy[6], Felipe I. Martínez[7,8] and Nicolás Montalva[9,10]*

[1]Criminal Justice Research Center, The Pennsylvania State University, University Park, PA, United States, [2]Comunidad Autónoma Diaguita Mapochogasta, Santiago, Chile, [3]Department of Human Genetics, University of Chicago, Chicago, IL, United States, [4]Department of Veterinary and Biomedical Sciences, College of Agricultural Sciences, The Pennsylvania State University, University Park, PA, United States, [5]The Huck Institutes of the Life Science, The Pennsylvania State University, University Park, PA, United States, [6]Comunidad Kawésqar, Puerto Edén, Chile, [7]School of Anthropology, Faculty of Social Sciences, Pontificia Universidad Católica, Santiago, Chile, [8]Center for Intercultural and Indigenous Research (CIIR), Santiago, Chile, [9]Society and Health Research Center, Universidad Mayor, Santiago, Chile, [10]School of Public Health, Universidad Mayor, Santiago, Chile

Keywords: genomics, mestizaje, Latin America, Indigenous, identity, research ethics, Chile, ancestry

OPEN ACCESS

Edited by:
Arslan A. Zaidi, University of Pennsylvania, United States

Reviewed by:
Tábita Hünemeier, University of São Paulo, Brazil

****Correspondence:***
Constanza P. Silva
cjs6251@psu.edu
Constanza de la Fuente Castro
cdelafuente@uchicago.edu
Tomás González Zarzar
tug156@psu.edu

Specialty section:
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

Received: *17 November 2021*
Accepted: *03 February 2022*
Published: *03 March 2022*

Citation:
Silva CP, de la Fuente Castro C, González Zarzar T, Raghavan M, Tonko-Huenucoy A, Martínez FI and Montalva N (2022) The Articulation of Genomics, Mestizaje, and Indigenous Identities in Chile: A Case Study of the Social Implications of Genomic Research in Light of Current Research Practices. Front. Genet. 13:817318. doi: 10.3389/fgene.2022.817318

1 INTRODUCTION

Genomic research has contributed significantly to our understanding of present-day human biological diversity, health, and disease. However, at the same time, genomic research has historically excluded marginalized groups. In the past decades, the increased access to genomic technologies and data has been paired with efforts to improve sampling diversity by including African descendants and Indigenous peoples. The rationale of these efforts is that disparities in research participation can potentially lead to inequalities in the benefits derived from genomic research (Lee, 2021). Nonetheless, this diversification push has had its pitfalls without clear protocols that improve protection to participants' DNA data access, data use, and intellectual property against commodification (Fox, 2020). Moreover, Indigenous participation in research continues to be framed in colonial power structures, which are often masked in "reciprocity" and "justice" undermining Indigenous peoples' sovereignty, self-determination, and governance (Tsosie et al., 2021).

In several Latin American countries, genomic research has advanced two apparently contradictory discourses framed within the *mestizo* rhetoric: *mestizaje*, perceived as genetic homogenization, and Indigenous "purism", understood as the existence of groups with unmixed genetic ancestry (TallBear, 2013b). These discourses entwine the ideas of genomic ancestry and identity and are perpetuated in academia, either by contributing to the erasure of Indigenous identities under the *Mestizo/Hispanic/Latino* umbrella or by the fetishization of indigeneity by using genetic categories based on a racial logic. The articulation of genetic ancestry with the *mestizo* and Indigenous identities, and their particularities, have been discussed more extensively in the past decades in Brazil, Colombia, and Mexico (Simpson, 2000; Beltrán et al., 2014; Wade et al., 2014b; Kent et al., 2014). It has been proposed that biogeographical ancestries (e.g., European, African, and Amerindian, also called Native American) can evoke different ideas of ancestry, appearance, culture, class, region, and nation, mainly among those outside the genetic research field (Wade et al., 2014b). However, we highlight that the genomic research approaches and interpretations by many experts in

Latin America reflect an ongoing global phenomenon in science where peoples' cultural histories and gene histories are entangled (TallBear, 2013a). This phenomenon of racialized thinking in genomics is embodied in the concept of nation or "genetic citizenships" (Wade et al., 2014b; Wade et al., 2015), which contributes to further stigmatization of historically discriminated populations.

This opinion contributes to this discussion by focusing on two main points. First, we discuss how genomic research opportunistically benefits from the two allegedly contradictory discourses present in the *mestizo* rhetoric; a single *mestizo* national identity based on genetic admixture and Indigenous "purism" based on the articulation of genetic diversity and ethnicity. Second, we situate this problem in the underexplored Chilean sociopolitical context and suggest strategies to improve Indigenous communities' agency in research settings, contributing to future guidelines on genomic research.

2 GENOMICS AND THE LATIN AMERICAN CONTEXT

2.1 Genomic Research and the *Mestizaje* Rhetoric

The invention of Latin America as a geopolitical entity had the aim to fulfill the promise of a "civilization" distanced from the Old World by creating a new cohesive identity (Torres Martínez, 2016). A critical component of this civilizing project was the *mestizaje* (or *mestiçagem* in Portuguese), defined as the admixture of different cultures and racial groups (Wade, 2003). For the Mexican philosopher José Vasconcelos, the *mestizo* represented a "cosmic race" that combined the virtues of Indigenous and Europeans by constituting *"the moral and material basis for the union of all men into a fifth universal race, the fruit of all the previous ones and amelioration of everything past."* (Vasconcelos and Sánchez, 1966). His ideas reflect the heart of the *mestizaje* as a racial project and building block for the construction of several Latin American national identities; rhetoric that culturally and racially homogenizes populations by erasing the Indigenous by either racial amalgamation or replacement (Tuck and Yang, 2012; Telles and Bailey, 2013). The *mestizaje* rhetoric has mutated in conjunction with the particular sociocultural history of each country, reflected in specific policies and institutions. Furthermore, it has also changed with the development of biotechnologies, including multidimensional ethno-racial concepts such as phenotype classifications, self-identification (Paredes, 2018), and now, DNA ancestry, showing how the implementation of these technologies are far from being neutral (Wade et al., 2014b).

Despite the efforts of scientists to distance themselves from racial categories, the scientific literature continues to use terms such as "DNA ancestry" similarly to racial categories (i.e., continental ancestries), which today are a component of the contemporary *mestizaje* rhetoric. Although the study of genetic variation can provide insights into the relatedness and migration histories of a person's ancestors, these are not equal to cultural identity or belonging (TallBear, 2013a; Roth et al., 2020). Thus, by positioning DNA as an essentialist marker of shared identity, there is a danger of equating genetic histories with cultural/ethnic identities (Simpson, 2000), a consequence that is permeating the *mestizaje* rhetoric.

In Latin America, the vast majority of human genomic projects have focused on estimating degrees of genetic admixture within nations (Acuña et al., 2000; Wade et al., 2014b; Eyheramendy et al., 2015; Homburger et al., 2015; Adhikari et al., 2016; Berrios, 2016). Such focus aimed at demonstrating that racial categories are useless and that genetic ancestry estimation would be the only deracialized approach to understand genetically mixed societies (Pena, 2000; Wade et al., 2014a; Kent et al., 2015; Mostrador, 2019). Some of these projects not only fall into a genetic fetishism to elucidate historical problematics (TallBear, 2013b), but also reinforce the discourse of a single mixed entity that cannot be differentiated (Kent et al., 2014), endorsing the *mestizo* rhetoric and nationhood (Séguin et al., 2008), yet avoiding any discussion about race (Rodríguez Mega, 2021). In addition, genomic studies have also focused on the genomic articulation of indigeneity or the essentialization of ethnicity into *imagined genetic communities* (Simpson, 2000) or discrete genetic clusters. This has been achieved by differentiating the *mestizo* from the Indigenous or by isolating one or several Indigenous genetic components (Wang et al., 2008; Verdugo et al., 2020).

We observe how genomic research in Latin America opportunistically profits from two apparently contradictory discourses: differentiation and homogeneity, both embedded in the *mestizaje* rhetoric under a scientific rationale. This rationale can achieve truthful results as it supposedly excludes the influence of cultural and social preconceptions (Wade et al., 2014b). However, scientists hold ethical, moral, and political positions that guide their research questions, objectives, approaches, results, and interpretations. The two assumptions ingrained in the *mestizaje* rhetoric are often the starting point of most research efforts in Latin America, such as the case of Chile. On the one hand, the *mestizaje* rhetoric treats genetic admixture as a *continuum*, where the degree of mestizaje can establish, allegedly, that some people are more Indigenous than others. This genetic admixture ladder is articulated with the attribute of indigeneity and, thus, privileges genetic ancestry as a proxy to define Indigenous populations (TallBear, 2013a; Walker et al., 2016). On the other hand, genetic admixture supports the *mestizo* rhetoric of unifying and homogenizing populations under a single national identity (Pena, 2000; Kent et al., 2015; Alpaslan-Roodenberg et al., 2021). Under this logic, the Indigenous/European admixture represented in the *mestizo* signifies the "genetic dissolution" of the original pre-Hispanic Indigenous, foregrounding the idea that either contemporary Indigenous peoples do not exist or are less Indigenous than their ancestors (Tuck and Yang, 2012). This conceptualization can have detrimental consequences in research and sovereignty for Indigenous peoples. For example, a recent publication on ancient DNA suggests that Indigenous heritage is embedded in the *mestizo* national identity of most Latin American countries (Alpaslan-Roodenberg et al., 2021). Therefore, the

implementation of similar research standards used in the US for Indigenous engagement and consultation could be counterproductive, and thus Indigenous consultation is not needed (Alpaslan-Roodenberg et al., 2021). However, the authors assume the total integration of Indigenous identities into the *mestizo* national identity, questioning the presence of Indigenous voices and disregarding sociocultural processes in these regions. Moreover, it represents a convenient assumption maintaining the status quo of research practices in Latin America and profiting from the lack of legal protection.

2.2 Indigenous Peoples and Genomic Research in Chile

In Chile, around 12.8% of the total population self-identifies as Indigenous. However, none of the 11 Indigenous groups (Mapuche, Aymara, Diaguita, Lickanantay, Quechua, Rapa Nui, Colla, Kawésqar, Chango, Yagán, and Selk'nam) are constitutionally recognized, meaning that collective and territorial rights, sovereignty, and self-determination are not guaranteed by the Chilean State (CIPERChile, 2019). Thus, there is an increased legal vulnerability of Indigenous peoples in Chile, compared to other Latin American countries (Fuentes et al., 2017). This situation is expected to change with the direct participation of Indigenous representatives in the current constituent assembly (Fuentes, 2021).

Chilean genomic research has followed the international trends, prioritizing the genetic characterization of the national *mestizo* population (Ruiz-Linares et al., 2014; Berrios, 2016; Paschetta et al., 2021), but also articulating genetic ancestry and ethnic identity by exploring the "origins" (i.e., ethnogenesis) of these populations (Acuña et al., 2000; Fuentes et al., 2014; Rothhammer et al., 2017; Verdugo et al., 2020). Furthermore, genomic research in Chile has also aimed at identifying informative ancestry markers to characterize specific clusters or ethnic groups (e.g., Mapuche and Aymara), in some cases linking them to diseases (Andia et al., 2008; Bermejo et al., 2017; Díaz-Peña et al., 2020; Jackson et al., 2021; Koenigstein et al., 2021). These concepts of genetic admixture, *mestizo*, or Indigenous DNA are continuously being tossed into the wider society and become part of popular discourses (Simpson, 2000), creating societal narratives about scientifically identifying who are the "real" Indigenous and a way to differentiate who belongs to a specific ethnic group.

Troubling narratives have been derived from genomic research on Indigenous peoples, such as the existence of a "Diaguita DNA" to be studied (ChileGenómico, 2019) and the genetic origin of the Chilean *mestizo* (Berrios, 2016), which serve to essentialize identities based on genetic categories. Genetic groupings can be variable and arbitrary across studies and go as far as labeling physical traits employing ethnonyms, all in a manner that worryingly brings race-based categorization to mind. In addition, studies about the correlation of disease biomarkers with Mapuche ancestry wrongly equate genetic ancestry to ethnicity as the cause of higher gallbladder cancer risk (Bermejo et al., 2017; Jackson et al., 2021). While genetic factors underlie disease susceptibilities, such singular narratives of Indigenous genetic ancestry percentage impacting disease risk drive us away from discussions regarding sociocultural (e.g., diet), socioeconomic, and geographical (rural vs. urban) factors, which could have a broader impact on prevention and health equity policies. Thus, although these research examples preach to be a first step to address racial health disparities, in reality, they evoke race-like categories (Wade et al., 2014b) that further stigmatize Indigenous peoples as genetically different from Chileans/Latinos. However, the bioethics committees at Chilean academic institutions still have to further develop specific protocols to incorporate Indigenous voices in conversations on ethical sampling procedures, informed consent, data privacy, result interpretations, and science communication. Although there is a need to create legal, regulatory, and normative instruments appropriate to the current challenges of genomic research that guarantee Indigenous communities' participation and protection of their genetic data, these changes will take time. We believe that, in the meantime, the alternative path is to empower Indigenous communities in research settings. This alternative gathers the sociopolitical responsibility for reparation that we, researchers, have in the face of a persistent history of bad practices.

3 STRATEGIES FOR MORE ETHICAL AND EQUITABLE RESEARCH

Often, research has been done on Indigenous peoples, instead of for, with, or by them (Dalton, 2002). Further, in territories where the government does not constitutionally recognize Indigenous peoples, the improvement of research practices in academia needs to start by empowering Indigenous communities in research settings instead of relying on the goodwill of the researchers or expecting academic institutions to make amends. There are great international examples of inclusive research by Indigenous researchers from the United States, Canada, Australia, and Aotearoa on community-based research, mentoring, and mechanisms to empower Indigenous peoples in Western research settings (Claw et al., 2018; Tsosie and Claw, 2020). In order to implement some of Claw et al., 2018 guidelines in the Chilean context, where this discussion is just beginning, we consider a need for a radical shift in how research is conducted. Therefore, we propose to: first, establish long-term partnerships between researchers and Indigenous communities with bidirectional educational purposes (Tsosie and Claw, 2020). This approach will allow the integration of cultural perspectives into research, which has the advantages of creating better-informed, ethical, culturally appropriate, and respectful science (Claw et al., 2018; Begay et al., 2020). Second, academic institutions and researchers should advocate and support educational, mentorship, and training opportunities for Indigenous peoples as researchers. For example, in the US, Canada, Australia, and Aotearoa, researchers have developed the *Summer internship for*

INdigenous peoples in Genomics (SING) workshop to discuss the uses, misuses, and limitations of genomics as a tool for Indigenous peoples' communities. Further, the long-term aim is to propel Indigenous peoples in science research, leadership, and teaching careers at all levels, making genomic research by and for Indigenous peoples. We expect these initial steps to promote a shift in the current research ethos in the region by improving research practices, scientific training, and moving towards community-based collaborative practices that support Indigenous interests and concerns. Many other areas still need to be improved regarding data privacy, data ownership, and research infrastructure within communities. Our suggestions are a first step for paving the path towards more ethical and beneficial research with Indigenous communities.

4 CONCLUSION

At the heart of all the above lies a somewhat urgent need to implement mechanisms in Latin America that can secure positive engagement with genomics while countering misuses of and misinformation from it. We believe that establishing novel collaborative mechanisms between academia and Indigenous groups can introduce researchers to knowledge that recognizes forms of kinship, relatedness, ancestry, and heritage that are not reliant on DNA; forms of knowledge that recognize the different places that people's histories and gene histories occupy. This collaborative approach is required to debunk myths about genetics and the *mestizo* rhetoric at large.

AUTHOR'S NOTE

Spanish version of this article is available in **Supplementary Material S1**.

AUTHOR CONTRIBUTIONS

CS gathered the literature data and drafted the manuscript. CS, CDC, and TGZ wrote the manuscript. CS, CDC, TGZ, MR, AT-H, FM, and NM conceptualized the idea of this manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

The authors acknowledge the support from CIIR/Fondap 15110006.

ACKNOWLEDGMENTS

The authors thank the University of Chicago Provost's Global Faculty Awards: Latin America.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.817318/full#supplementary-material

REFERENCES

Acuña, M., Llop, E., and Rothhammer, F. (2000). Genetic Composition of Chilean Population: Rural Communities of Elqui, Limari and Choapa Valleys. *Rev. Med. Chil* 128, 593–600. doi:10.4067/s0034-98872000000600004

Adhikari, K., Mendoza-Revilla, J., Chacón-Duque, J. C., Fuentes-Guajardo, M., and Ruiz-Linares, A. (2016). Admixture in Latin America. *Curr. Opin. Genet. Dev.* 41, 106–114. doi:10.1016/j.gde.2016.09.003

Alpaslan-Roodenberg, S., Anthony, D., Babiker, H., Bánffy, E., Booth, T., Capone, P., et al. (2021). Ethics of DNA Research on Human Remains: Five Globally Applicable Guidelines. *Nature* 599, 41–46. doi:10.1038/s41586-021-04008-x

Andia, M. E., Hsing, A. W., Andreotti, G., and Ferreccio, C. (2008). Geographic Variation of Gallbladder Cancer Mortality and Risk Factors in Chile: A Population-based Ecologic Study. *Int. J. Cancer* 123, 1411–1416. doi:10.1002/ijc.23662

Begay, R. L., Garrison, N. A., Sage, F., Bauer, M., Knoki-Wilson, U., Begay, D. H., et al. (2019). Weaving the Strands of Life (Iiná Bitł'ool): History of Genetic Research Involving Navajo People. *Hum. Biol.* 91, 189–208. doi:10.13110/humanbiology.91.3.04

Beltrán, C. L., Deister, V. G., and Sandoval, M. R. (2020). "3 Negotiating the Mexican Mestizo: On the Possibility of a National Genomics," in *Mestizo Genomics: Race Mixture, Nation, and Science in Latin America* (Durham: Duke University Press), 85–106. doi:10.1515/9780822376729-006

Bermejo, J. L., Boekstegers, F., González Silos, R., Marcelain, K., Baez Benavides, P., Barahona Ponce, C., et al. (2017). Subtypes of Native American Ancestry and Leading Causes of Death: Mapuche Ancestry-specific Associations with Gallbladder Cancer Risk in Chile. *Plos Genet.* 13, e1006756. doi:10.1371/journal.pgen.1006756

Berrios, S. (2016). *El ADN de los chilenos y sus orígenes genéticos*. Santiago: Editorial Universitaria de Chile.

ChileGenómico (2019). *Chilegenomico se adjudica fondos para investigar el ADN diaguita – Chilegenómico*. Available at: https://www.chilegenomico.cl/2019/08/28/fondos-and-diaguita/.

CIPERChile (2019). *Plurinacionalidad y reconocimiento de los pueblos: las demandas indígenas para la nueva Constitución*. Available at: https://www.ciperchile.cl/2019/11/14/plurinacionalidad-y-reconocimiento-de-los-pueblos-las-demandas-indigenas-para-la-nueva-constitucion/.

Claw, K. G., Anderson, M. Z., Anderson, M. Z., Begay, R. L., Tsosie, K. S., Fox, K., et al. (2018). A Framework for Enhancing Ethical Genomic Research with Indigenous Communities. *Nat. Commun.* 9, 2957. doi:10.1038/s41467-018-05188-3

Dalton, R. (2002). Tribe Blasts 'exploitation' of Blood Samples. *Nature* 420, 111. doi:10.1038/420111a

Díaz-Peña, R., Boekstegers, F., Silva, R. S., Jaime, S., Hosgood, H. D., Miravitlles, M., et al. (2020). Amerindian Ancestry Influences Genetic Susceptibility to Chronic Obstructive Pulmonary Disease. *Jpm* 10, 93. doi:10.3390/jpm10030093

Eyheramendy, S., Martinez, F. I., Manevy, F., Vial, C., and Repetto, G. M. (2015). Genetic Structure Characterization of Chileans Reflects Historical Immigration Patterns. *Nat. Commun.* 6, 6472. doi:10.1038/ncomms7472

Fox, K. (2020). The Illusion of Inclusion - the "All of Us" Research Program and Indigenous Peoples' DNA. *N. Engl. J. Med.* 383, 411–413. doi:10.1056/NEJMp1915987

Fuentes, C. (2021). *Pueblos originarios: entre el reconocimiento cultural y el reconocimiento político*.

Fuentes, C., Cea, de, M., Fuentes, C., and de Cea, M. (2017). Reconocimiento débil: derechos de pueblos indígenas en Chile. *pl* 25, 55–75. doi:10.18504/pl2549-003-2017

Fuentes, M., Pulgar, I., Gallo, C., Bortolini, M.-C., Canizales-Quinteros, S., Bedoya, G., et al. (2014). Geografía génica de Chile: Distribución regional de los aportes genéticos americanos, europeos y africanos. *Rev. Méd. Chile* 142, 281–289. doi:10.4067/S0034-98872014000300001

Homburger, J. R., Moreno-Estrada, A., Gignoux, C. R., Nelson, D., Sanchez, E., Ortiz-Tello, P., et al. (2015). Genomic Insights into the Ancestry and Demographic History of South America. *Plos Genet.* 11, e1005602. doi:10.1371/journal.pgen.1005602

Jackson, S. S., Van De Wyngard, V., Pfeiffer, R. M., Cook, P., Hildesheim, A., Pinto, L. A., et al. (2021). Inflammatory Profiles in Chilean Mapuche and Non-mapuche Women with Gallstones at Risk of Developing Gallbladder Cancer. *Sci. Rep.* 11, 3686. doi:10.1038/s41598-021-83300-2

Kent, M., García-Deister, V., López-Beltrán, C., Santos, R. V., Schwartz-Marín, E., and Wade, P. (2015). Building the Genomic Nation: 'Homo Brasilis' and the 'Genoma Mexicano' in Comparative Cultural Perspective. *Soc. Stud. Sci.* 45, 839–861. doi:10.1177/0306312715611262

Kent, M., Santos, R. V., and Wade, P. (2014). Negotiating Imagined Genetic Communities: Unity and Diversity in Brazilian Science and Society. *Am. Anthropologist* 116, 736–748. doi:10.1111/aman.12142

Koenigstein, F., Boekstegers, F., Wilson, J. F., Fuentes-Guajardo, M., Gonzalez-Jose, R., Bedoya, G., et al. (2021). Inbreeding, Native American Ancestry and Child Mortality: Linking Human Selection and Paediatric Medicine. *Hum. Mol. Genet.*, ddab302. doi:10.1093/hmg/ddab302

Lee, S. S.-J. (2021). Obligations of the "Gift": Reciprocity and Responsibility in Precision Medicine. *Am. J. Bioeth.* 21, 57–66. doi:10.1080/15265161.2020.1851813

Mostrador, E. (2019). Buscan secuenciar ADN de los diaguitas, habitantes precolombinos de Chile. Available at: https://www.elmostrador.cl/dia/2019/03/10/buscan-secuenciar-adn-de-los-diaguitas-habitantes-precolombinos-de-chile/.

Paredes, C. L. (2018). Multidimensional Ethno-Racial Status in Contexts of Mestizaje: Ethno-Racial Stratification in Contemporary Peru. *Socius* 4, 237802311876200. doi:10.1177/2378023118762002

Paschetta, C., de Azevedo, S., Ramallo, V., Cintas, C., Pérez, O., Navarro, P., et al. (2021). The Impact of Socioeconomic and Phenotypic Traits on Self-Perception of Ethnicity in Latin America. *Sci. Rep.* 11, 12617. doi:10.1038/s41598-021-92061-x

Pena, S. (2000). Revista Ciência Hoje - Retrato Molecular Do Brasil. Available at: https://issuu.com/biohoffmann/docs/ci_ncia_hoje__v.27__n.159__abr__2000_-_retrato_mol.

Rodríguez Mega, E. (2021). How the Mixed-Race Mestizo Myth Warped Science in Latin America. *Nature* 600, 374–378. doi:10.1038/d41586-021-03622-z

Roth, W. D., Yaylacı, Ş., Jaffe, K., and Richardson, L. (2020). Do genetic Ancestry Tests Increase Racial Essentialism? Findings from a Randomized Controlled Trial. *PLoS ONE* 15, e0227399. doi:10.1371/journal.pone.0227399

Rothhammer, F., Puddu, G., Fuentes-Guajardo, M., Rothhammer, F., Puddu, G., and Fuentes-Guajardo, M. (2017). ¿Puede El Adn Mitocondrial Proporcionar Información Sobre La Etnogénesis De Los Pueblos Originarios Chilenos? *Chungará (Arica)* 49, 635–642. doi:10.4067/S0717-73562017005000028

Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., et al. (2014). Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals. *Plos Genet.* 10, e1004572. doi:10.1371/journal.pgen.1004572

Séguin, B., Hardy, B.-J., Singer, P. A., and Daar, A. S. (2008). Genomics, Public Health and Developing Countries: The Case of the Mexican National Institute of Genomic Medicine (INMEGEN). *Nat. Rev. Genet.* 9, S5–S9. doi:10.1038/nrg2442

Simpson, B. (2000). Imagined Genetic Communities: Ethnicity and Essentialism in the Twenty-First century. *Anthropol. Today* 16, 3–6. doi:10.1111/1467-8322.00023

TallBear, K. (2013b). *Native American DNA: Tribal Belonging and the False Promise of Genetic Science*. Minneapolis: University of Minnesota Press.

TallBear, K. (2013a). Genomic Articulations of Indigeneity. *Soc. Stud. Sci.* 43, 509–533. doi:10.1177/0306312713483893

Telles, E., and Bailey, S. (2013). Understanding Latin American Beliefs about Racial Inequality. *Am. J. Sociol.* 118, 1559–1595. doi:10.1086/670268

Torres Martínez, R. (2016). Sobre el concepto de América Latina¿Invención francesa? *Cahiers D'études Romanes. Revue Du CAER*, 89–98. doi:10.4000/etudesromanes.5141

Tsosie, K. S., and Claw, K. G. (2020). Indigenizing Science and Reasserting Indigeneity in Research. *Hum. Biol.* 91, 137–140. doi:10.13110/humanbiology.91.3.02

Tsosie, K. S., Yracheta, J. M., Kolopenuk, J. A., and Geary, J. (2021). We Have "Gifted" Enough: Indigenous Genomic Data Sovereignty in Precision Medicine. *Am. J. Bioeth.* 21, 72–75. doi:10.1080/15265161.2021.1891347

Tuck, E., and Yang, K. W. (2012). Decolonization Is Not a Metaphor. *Decolonization: Indigeneity, Edu. Soc.* 1, 1–40.

Vasconcelos, J., and Sánchez, L. A. (1966). *La Raza Cósmica: Misión de La Raza Iberoamericana*. Madrid: Aguilar.

Verdugo, R. A., Di Genova, A., Herrera, L., Moraga, M., Acuña, M., Berríos, S., et al. (2020). Development of a Small Panel of SNPs to Infer Ancestry in Chileans that Distinguishes Aymara and Mapuche Components. *Biol. Res.* 53, 15. doi:10.1186/s40659-020-00284-5

Wade, P., Deister, V. G., Kent, M., Olarte Sierra, M. F., and del Castillo Hernández, A. D. (2014a). Nation and the Absent Presence of Race in Latin American Genomics. *Curr. Anthropol.* 55, 497–522. doi:10.1086/677945

Wade, P., López Beltrán, C., Restrepo, E., and Ventura Santos, R. (2014b). *Mestizo Genomics: Race Mixture, Nation, and Science in Latin America*. Durham: Duke University Press.

Wade, P., López-Beltrán, C., Restrepo, E., and Santos, R. V. (2015). Genomic Research, Publics and Experts in Latin America: Nation, Race and Body. *Soc. Stud. Sci.* 45, 775–796. doi:10.1177/0306312715623108

Wade, P. (2003). Repensando El Mestizaje. *rev.colomb.antropol.* 39, 273–296. doi:10.22380/2539472x.1243

Walker, A., Egan, B., and Nicholas, G. (2016). DNA and Indigeneity Public Symposium & Workshop. Available at: https://www.sfu.ca/ipinch/events/ipinch-events/public-symposium-dna-and-indigeneity/.

Wang, S., Ray, N., Rojas, W., Parra, M. V., Bedoya, G., Gallo, C., et al. (2008). Geographic Patterns of Genome Admixture in Latin American Mestizos. *Plos Genet.* 4, e1000037. doi:10.1371/journal.pgen.1000037

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

PERSPECTIVE
published: 21 March 2022
doi: 10.3389/fgene.2022.823309

Using Indigenous Standards to Implement the CARE Principles: Setting Expectations through Tribal Research Codes

Stephanie Russo Carroll[1,2], *Ibrahim Garba*[1,2], *Rebecca Plevel*[2,3,4], *Desi Small-Rodriguez*[5,6], *Vanessa Y. Hiratsuka*[7], *Maui Hudson*[8] *and Nanibaa' A. Garrison*[9,10,11]*

[1]*Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ, United States,* [2]*Native Nations Institute, Udall Center for Studies in Public Policy, University of Arizona, Tucson, AZ, United States,* [3]*Library and Information Sciences, School of Information, University of Arizona, Tucson, AZ, United States,* [4]*Law Library, School of Law, University of South Carolina, Columbia, SC, United States,* [5]*Department of Sociology, College of Social Sciences, University of California, Los Angeles, Los Angeles, CA, United States,* [6]*American Indian Studies Program, College of Social Sciences, University of California, Los Angeles, Los Angeles, CA, United States,* [7]*Center for Human Development, College of Health, University of Alaska Anchorage, Anchorage, AK, United States,* [8]*Te Kotahi Research Institute, University of Waikato, Hamilton, New Zealand,* [9]*Institute for Society and Genetics, University of California, Los Angeles, Los Angeles, CA, United States,* [10]*Institute for Precision Health, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States,* [11]*Division of General Internal Medicine and Health Services Research, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States*

OPEN ACCESS

Edited by:
Tina Lasisi,
The Pennsylvania State University (PSU), United States

Reviewed by:
Nchangwi S. Munung,
University of Cape Town, South Africa
John C. Burnett,
Beckman Research Institute, City of Hope, United States
Rebecca Pollet,
Vassar College, United States

***Correspondence:**
Nanibaa' A. Garrison
nanibaa@socgen.ucla.edu

Specialty section:
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

Received: *27 November 2021*
Accepted: *21 February 2022*
Published: *21 March 2022*

Citation:
Carroll SR, Garba I, Plevel R, Small-Rodriguez D, Hiratsuka VY, Hudson M and Garrison NA (2022) Using Indigenous Standards to Implement the CARE Principles: Setting Expectations through Tribal Research Codes. Front. Genet. 13:823309. doi: 10.3389/fgene.2022.823309

Biomedical data are now organized in large-scale databases allowing researchers worldwide to access and utilize the data for new projects. As new technologies generate even larger amounts of data, data governance and data management are becoming pressing challenges. The FAIR principles (Findable, Accessible, Interoperable, and Reusable) were developed to facilitate data sharing. However, the Indigenous Data Sovereignty movement advocates for greater Indigenous control and oversight in order to share data on Indigenous Peoples' terms. This is especially true in the context of genetic research where Indigenous Peoples historically have been unethically exploited in the name of science. This article outlines the relationship between sovereignty and ethics in the context of data to describe the collective rights that Indigenous Peoples assert to increase control over their biomedical data. Then drawing on the CARE Principles for Indigenous Data Governance (Collective benefit, Authority to control, Responsibility, and Ethics), we explore how standards already set by Native nations in the United States, such as tribal research codes, provide direction for implementation of the CARE Principles to complement FAIR. A broader approach to policy and procedure regarding tribal participation in biomedical research is required and we make recommendations for tribes, institutions, and ethical practice.

Keywords: genetic research, Indigenous, data sovereignty, data governance, CARE principles

INTRODUCTION

As technological advances have generated immense amounts of biomedical data, the Indigenous Data Sovereignty (IDSov) movement has emerged to exert stronger control and oversight over data generated from Indigenous Peoples. Once subject to localized systems of management, biomedical data are now organized and stored in large-scale databases, allowing researchers worldwide to access and utilize data for new analyses. The governance of large-scale databases, many of which adopt broad data sharing models, often stands in contrast with stricter mechanisms of protection and relationships of trust that facilitated the original data collection. This disconnect is clearly evident in the case of Indigenous communities who have often challenged the extractive nature of genetic research (Boyer et al., 2007; Shaw et al., 2013; Trinidad et al., 2015; Haring et al., 2018; Chadwick et al., 2019; Dirks et al., 2019). We support the call for more open, inclusive, and equitable participation in research and innovation to resolve the tension between openness and innovation, on the one hand, and Indigenous rights and interests, on the other. This is a tension that pervades the current discourse on genetic diversity (Hudson et al., 2020; Welch et al., 2021).

Historically, biomedical data may not have been collected or utilized in ways that align with community rights and interests. The results are research with little or no benefit to the communities from which the data originated, potential biases in data interpretation, dwindling participation in genetics and genomics research, and limited oversight by the people from whom the data are collected (Garrison et al., 2019a). These negative experiences compound as biomedical and data futures move towards big data and large-scale biobanking. At the same time, the resurgence of Indigenous self-determination and the advancement of IDSov prompts a reexamination of data governance (Kukutai and Taylor 2016a; Garrison et al., 2019a; Carroll et al., 2020; Hudson et al., 2020; Walter et al., 2021). At a fundamental level, IDSov articulates the rights of Indigenous Peoples and nations to govern the collection, application, and use of data about their peoples, communities, lands, and resources (Kukutai and Taylor, 2016b).

This article outlines the relationship between sovereignty and ethics in the context of data to describe the collective rights that Indigenous Peoples assert to increase control over their biomedical data. Then drawing on the CARE Principles for Indigenous Data Governance (Collective benefit, Authority to control, Responsibility, and Ethics), we explore how standards already set by Native nations, such as tribal research codes, provide direction for implementing the CARE Principles. We close with recommendations for using tribal codes, laws, policy documents, and protocols to operationalize the CARE Principles as a way to spur translational genetics research that benefits Native nations, as well as rural and urban Indigenous communities.

Indigenous Peoples and Data

For the purposes of this paper, we define Indigenous Peoples in the US as American Indian, Alaska Native, Native Hawaiian, and other communities who are indigenous to the US and its territories. We will use Native nations and tribes interchangeably to refer to tribal nations in the US. The federal government recognizes 574 tribes in the US as sovereign nations with their own legal and political structures to govern their citizens and homelands (Department of the Interior, 2021). In addition, many other Indigenous Peoples exert sovereignty as state-recognized (National Conference of State Legislatures, 2019) or un-recognized nations, including those in the state of Hawai'i and US territories. Sovereignty refers to the collective powers of a nation, such as the power to grant access to the population or to negotiate treaties between nations. As sovereign nations, tribes have the power to govern via their own structures, determine their own citizenship, and regulate tribal business (Duthu, 2008).

Indigenous Peoples have always been "researchers," demonstrated by their collecting, analyzing, and managing data for decision-making, knowledge transfer, and other uses. Historical and ongoing colonialism disrupted, co-opted, and suppressed Indigenous research methodologies and methods (Smith, 2012). Indigenous data, whether born digital or not, include information, knowledge, specimens, and belongings about Indigenous Peoples to which they relate at both the individual and collective levels (Carroll et al., 2020; Rainie et al., 2019; Lovett et al., 2019). IDSov returns authority over data about Indigenous nations and their citizens, communities, and resources (wherever they may be located) back to the tribes from whom the data derive (Kukutai and Taylor, 2016b). Indigenous Data Governance (IDGov) enables tribal ways of knowing and doing to guide Indigenous decision-making; it is a practical expression of IDSov (Rainie et al., 2017; Maiam nayri Wingara, 2018).

Increasingly over the past 50 years, tribes in the US have developed policies and procedures for the oversight of research within their nations' physical jurisdiction and beyond tribal lands. Other Native nations rely on tribal colleges, tribal organizations, or the Indian Health Service to provide research oversight on their behalf (Around Him et al., 2019). Federally-recognized tribes are in the strongest legal position to assert authority over their data (Tsosie, 2019). Non-federally-recognized tribes and Indigenous Peoples worldwide experience numerous issues in exercising rights over their data that may be different from federally-recognized tribes (Kukutai and Taylor, 2016a; Walter et al., 2021). However, we posit that learnings from federally-recognized tribes' codes can broadly benefit Indigenous Peoples as they implement laws, policies, and practices to govern their data and research.

IDGov and tribal research governance complement one another: some data are research data that are subject to both data governance and research governance. Thus, Indigenous research governance becomes a mechanism for enhancing IDGov as tribes assert IDSov.

Indigenous Peoples' Increased Oversight of Biomedical Research

IDSov requires heightened consideration in projects that evoke a government-to-government relationship, such as federally funded projects that seek to recruit large numbers of Indigenous Peoples nationwide. In these cases, strong relationships and effective data governance systems at the tribal level are paramount for ensuring equitable participation in federally funded research and culturally rigorous results. At the same time, non-tribal institutional policies and practices must also evolve to promote and protect the sovereign rights and interests of Indigenous Peoples.

American Indian and Alaska Native populations are not simply ethnic or racial groups, nor are they vulnerable or "special" populations. Tribes maintain a unique political status and confer citizenship just like other nation states. Tribal citizenship persists regardless of residence on or off tribal lands. Also called tribal enrollment, tribal citizenship is not the same as self-identification nor is it the same as genetic ancestry (Tallbear, 2013). Tribal citizenship is a political designation similar to US citizenship. This political designation is the foundation for IDSov. Yet, the inclusion of Indigenous people off tribal lands challenges the reach of tribal oversight of research over enrolled tribal citizens. Approximately 78% of self-identified American Indian and Alaska Native individuals live off tribal lands, and approximately 60% primarily reside in urban areas (Norris et al., 2012). For Indigenous people living off tribal lands, questions arise regarding how tribes will govern information about them when data are collected and reside outside the jurisdictional boundaries of the tribal nation. Additional questions include how other institutions, such as intertribal non-profit organizations and universities, will steward and protect data about Indigenous Peoples and individuals.

The recognition of IDSov by federal agencies and large repositories funded by organizations like the National Institutes of Health (NIH) is an important first step. The use of already existing tribal expectations delineated in reports, policies, and practices are important next steps to align federal programs with tribal rights and expectations via IDSov (Tribal Collaboration Working Group, 2018). In late 2019, the National Congress of American Indians (NCAI) (National Congress of American Indians, 2019a) asserted that even in the absence of formal tribal approval processes, researchers must establish a process to obtain approval that allows for tribal oversight of tribal data. Furthermore, the NCAI membership passed Resolution ABQ-19-061 that "calls on NIH to consult with tribal nations, provide a process for tribal nations to have oversight over any data and biospecimens from their tribal citizens, and restrict use of data associated with tribal nations until tribal oversight is in place" (National Congress of American Indians, 2019b). This resulted in developing a formal tribal consultation process (National Congress of American Indians, 2021; Haozous et al., 2021).

Tribal concerns about data use and data sharing have generated many discussions in federal agencies, universities, professional societies, and Indigenous communities. To build ethical university-tribal partnerships, it is necessary to recognize tribes as sovereign nations, acknowledge tribal intellectual property, and respect tribal data sharing preferences (James et al., 2014). Indigenous individuals' concerns about privacy and confidentiality also extend to promotion of tribal rights to control data and protection of collective tribal confidentiality and privacy in data and research (Taitingfong et al., 2020). In interviews with Indigenous leaders, scholars, and tribal research review members, support for tribal oversight of data is seen as a viable solution to the challenges of data access, management, and sharing (Garrison et al., 2019b). Given the history of exploitative research with tribal communities, the ability of tribes to review inaccurate, harmful, or stigmatizing information before publication or distribution is crucial both to preventing the misuse of their data and to supporting sound scientific practice (Garrison et al., 2019a). This is increasingly important as biomedical and genomics research moves toward broad data sharing policies.

Indigenous data oversight has increased in response to support of broad data sharing by funders and scientists. The NIH Genomic Data Sharing policy requires federally-funded investigators to deposit de-identified data into federal databases to promote secondary analyses (National Institutes of Health, 2014). However, the policy allows a data sharing exception that recognizes some tribal laws may not permit broad data sharing (Hiratsuka et al., 2020). Some tribal laws and policies dictate that all data generated from a research study is property of the tribe and all data must be returned to the tribe at the conclusion of the study. A resulting concern about the data sharing policy is that the allowable exceptions are not clearly understood or recognized by all researchers, institutions, or journal editors. For example, some investigators who conduct research with Indigenous communities have been asked by journal editors to submit the data to federal databases, even when the agreement with the tribe is not to share data.

IMPLEMENTATION OF CARE PRINCIPLES GUIDED BY TRIBAL OVERSIGHT

The current structures that are in place for federal biomedical data governance, in particular the Common Rule (Office of Human Research Protections, 2017), fail to align with the rights and interests of Indigenous nations and communities (Hudson et al., 2020). Rather than demanding that representatives of Indigenous communities participate in these existing governance structures, we argue for sovereign control—that is, Indigenous nations controlling ownership, governing storage, and dictating parameters for data use and reuse. We also promote policy innovations for other institutions that both adhere to tribal sovereignty and protect Indigenous people living off tribal lands or who self-identify as Indigenous (i.e., not tribally affiliated).

This section introduces the CARE Principles for Indigenous Data Governance as high-level guidance for enhancing IDSov in research and data governance. This section also examines the sovereign expectations that tribes set for researchers and institutions to

TABLE 1 | The CARE Principles for Indigenous Data Governance and sub-principles.

COLLECTIVE BENEFIT: Data ecosystems shall be designed and function in ways that enable Indigenous Peoples to derive benefit from the data
C1: For Inclusive Development and Innovation Governments and institutions must actively support the use and reuse of data by Indigenous nations and communities by facilitating the establishment of the foundations for Indigenous innovation, value generation, and the promotion of local self-determined development processes C2: For Improved Governance and Citizen Engagement Data enrich the planning, implementation, and evaluation processes that support the service and policy needs of Indigenous communities. Data also enable better engagement between citizens, institutions, and governments to improve decision-making. Ethical use of open data has the capacity to improve transparency and decision-making by providing Indigenous nations and communities with a better understanding of their peoples, territories, and resources. It similarly can provide greater insight into third-party policies and programs affecting Indigenous Peoples C3: For Equitable Outcomes Indigenous data are grounded in community values, which extend to society at large. Any value created from Indigenous data should benefit Indigenous communities in an equitable manner and contribute to Indigenous aspirations for wellbeing
AUTHORITY TO CONTROL: Indigenous Peoples' rights and interests in Indigenous data must be recognised and their authority to control such data be empowered. Indigenous data governance enables Indigenous Peoples and governing bodies to determine how Indigenous Peoples, as well as Indigenous lands, territories, resources, knowledges and geographical indicators, are represented and identified within data
A1: Recognizing Rights and Interests Indigenous Peoples have rights and interests in both Indigenous Knowledge and Indigenous data. Indigenous Peoples have collective and individual rights to free, prior, and informed consent in the collection and use of such data, including the development of data policies and protocols for collection A2: Data for Governance Indigenous Peoples have the right to data that are relevant to their world views and empower self-determination and effective self-governance. Indigenous data must be made available and accessible to Indigenous nations and communities in order to support Indigenous governance A3: Governance of Data Indigenous Peoples have the right to develop cultural governance protocols for Indigenous data and be active leaders in the stewardship of, and access to, Indigenous data especially in the context of Indigenous Knowledge
RESPONSIBILITY: Those working with Indigenous data have a responsibility to share how those data are used to support Indigenous Peoples' self-determination and collective benefit. Accountability requires meaningful and openly available evidence of these efforts and the benefits accruing to Indigenous Peoples
R1: For Positive Relationships Indigenous data use is unviable unless linked to relationships built on respect, reciprocity, trust, and mutual understanding, as defined by the Indigenous Peoples to whom those data relate. Those working with Indigenous data are responsible for ensuring that the creation, interpretation, and use of those data uphold, or are respectful of, the dignity of Indigenous nations and communities R2: For Expanding Capability and Capacity Use of Indigenous data invokes a reciprocal responsibility to enhance data literacy within Indigenous communities and to support the development of an Indigenous data workforce and digital infrastructure to enable the creation, collection, management, security, governance, and application of data R3: For Indigenous Languages and Worldviews Resources must be provided to generate data grounded in the languages, worldviews, and lived experiences (including values and principles) of Indigenous Peoples
ETHICS: Indigenous Peoples' rights and wellbeing should be the primary concern at all stages of the data life cycle and across the data ecosystem
E1: For Minimizing Harm and Maximizing Benefit Ethical data are data that do not stigmatize or portray Indigenous Peoples, cultures, or knowledges in terms of deficit. Ethical data are collected and used in ways that align with Indigenous ethical frameworks and with rights affirmed in UNDRIP. Assessing ethical benefits and harms should be done from the perspective of the Indigenous Peoples, nations, or communities to whom the data relate E2: For Justice Ethical processes address imbalances in power, resources, and how these affect the expression of Indigenous rights and human rights. Ethical processes must include representation from relevant Indigenous communities E3: For Future Use Data governance should take into account the potential future use and future harm based on ethical frameworks grounded in the values and principles of the relevant Indigenous community. Metadata should acknowledge the provenance and purpose and any limitations or obligations in secondary use inclusive of issues of consent

support Indigenous Peoples' efforts to reclaim control and oversight of data, including biospecimens.

The CARE Principles for Indigenous Data Governance

The CARE Principles define Collective benefit, Authority to control, Responsibility, and Ethics, and their relationship to engagement with and for secondary use of Indigenous data (Research Data Alliance Interest Group, 2019). The CARE Principles and the sub-principles (see **Table 1**) enhance and extend the 'FAIR Principles' for scientific data management (Findable, Accessible, Interoperable, Reusable; Wilkinson et al., 2016) by centering equity and ethics as core guiding principles alongside those set out by FAIR. The CARE Principles reflect the crucial role of data in advancing

TABLE 2 | The CARE Principles for Indigenous Data Governance: Tribal expectations that guide implementation.

Principle/Sub-principle	Quotes from Tribal Documents	Tribal Expectations
COLLECTIVE BENEFIT: Data ecosystems, including research life cycle, to be organized in ways open to collective Indigenous input and accessible for collective Indigenous benefit		
C1: For Inclusive Development and Innovation	Researchers shall provide for Tribal oversight of projects and report regularly to the Tribal Council and liaison department of project progress and results. Confederated Tribes of Siletz Indians, (2005) The tribe will only support community engaged research practices, which requires a high level of collaboration with Cherokee Nation (integrating the ideas of the tribal into the study) and must address Cherokee needs to benefit the citizens. Cherokee Nation, (2019b)	Project outcomes to align with tribal needs and tribal input to be incorporated into research process
C2: For Improved Governance and Citizen Engagement	Research should not be conducted until there has been full consultation with all potentially affected communities and individuals including all human research subjects, and each such community and individual have approved the research after full disclosure. Turtle Mountain Band of Chippewa Indians, (2014) Researchers are advised to budget funding...to provide adequate resources to cover community education and outreach efforts. Mohawk Nation of Akwesasne, (1996)	Obligation to engage, consult, and seek approval of both individuals and communities potentially affected by the research
C3: For Equitable Outcomes	Expected benefits of the proposed research, primary or secondary findings, including immediate and long range benefits to... the Nation; the Indian people generally; and society generally. Ho-Chunk Nation, (2005) Just compensation or fair return includes but is not limited to: obtaining copies of the research findings, authorship, co-authorship or acknowledgment, royalties, fair monetary compensation, copyright, patent, trademark. Mohawk Nation of Akwesasne, (1996)	Benefits may apply broadly but such benefits must have specific connections to tribal needs and priorities
AUTHORITY TO CONTROL: Recognition of Indigenous rights regarding research materials and data involve return of findings to community and control of uses outside tribal territory		
A1: Recognizing Rights and Interests	Principle of Prior Rights: This principle recognizes that Indigenous peoples, traditional societies, and local communities have prior, proprietary rights and interests over all air, land, and waterways, and the natural resources within their territories that they have traditionally inhabited or used, together with all knowledge and intellectual property and traditional resource rights associated with such resources and their use. Turtle Mountain Band of Chippewa Indians, (2014) This Code shall apply to all research (as defined elsewhere in this Code) conducted within the Nation's Territory, whether involving human subjects or not, and all research regarding materials wherever located as to which the Nation has a claim of intellectual, cultural or other ownership, legal or equitable. Ho-Chunk Nation, (2005)	Tribal claims to ownership of research materials and data, and expressions of prior Indigenous rights to lands, waterways, and natural resources
A2: Data for Governance	The process of developing community-based and culturally relevant research should directly include the tribe from the studies inception and supports a tribal agenda (plus whenever possible include local Native American investigators). Cherokee Nation, (2019b) At a minimum, the following information shall be provided by a Medical and Health Care applicant researcher ... (G) ... opportunity for the Community, Districts, and individuals, as appropriate to have periodic reports on the progress of the Medical Health Care Research and to comment on periodic and draft final reports. Gila River Indian Community, (2009)	Findings from research to be returned to the community to support governance and self-determination
A3: Governance of Data	Research information and data generated by and about Navajo individuals, communities, culture represent inalienable intellectual properties of the Navajo people and over which the Navajo Nation will provide oversight. Navajo Nation, (2002) This principle recognizes that the Tribe and any human research subjects, at its/their sole discretion, have the right to exclude from publication and/or to have kept confidential, any information including information concerning themselves, their health, or their culture, traditional knowledge, traditions, mythologies, or spiritual beliefs ... Three Affiliated Tribes (n.d.)	Tribal governments have right and responsibility to ensure research data used in ways consistent with community values, interests, and priorities

(Continued on following page)

TABLE 2 | (*Continued*) The CARE Principles for Indigenous Data Governance: Tribal expectations that guide implementation.

Principle/Sub-principle	Quotes from Tribal Documents	Tribal Expectations
RESPONSIBILITY: Researchers to respect Indigenous classifications, restrictions, and practices in relation to data and to advance community's capacity to manage own data by involving members in research activities		
R1: For Positive Relationships	This principle recognizes the necessity for researchers to respect the integrity, morality, and spirituality of the culture, traditions, and relationships of Tribal members with the world, and to avoid the imposition of external conceptions and standards. Turtle Mountain Band of Chippewa Indians, (2014) Cultural sensitivity training for the researchers as well as research awareness presentations on the Reservation will help develop a mutual understanding in conducting the research projects. Three Affiliated Tribes (n.d.)	Mutual understanding and respect crucial in engaging Indigenous data, especially those data considered sacred or culturally significant
R2: For Expanding Capability and Capacity	The Research Advisory Committee will help to ensure that the proposed research... empowers those involved through education, training and/or authorship. Mohawk Nation of Akwesasne, (1996) . . . Provisions for Native and local preference in employment in all phases of the project, including both on and off Reservation phases. White Earth Nation, (2018)	Researchers to strengthen community's ability to manage own data through training and employment opportunities in research projects
R3: For Indigenous Languages and Worldviews	Further, the Karuk Tribe asserts its age-old tradition of reserving domains of knowledge for rightful and culturally appropriate owners, as well as restricting access to this knowledge during certain chronological periods as dictated by time honored Karuk Law. Karuk Tribe, (2015) "Human Subject" means a living or nonliving individual (including human remains) about whom a researcher conducting research obtains information or data through interaction with the individual, involving physical procedures by which data are gathered (for example, blood draws), and/or manipulations of the subject or the subject's environment. Tohono O'odham Nation (2013)	Recognition and inclusion of Indigenous data norms and practices throughout research process
ETHICS: Obligation to minimize risks and maximize community benefits throughout research life cycle and also to strengthen Indigenous rights by addressing power and other imbalances		
E1: For Minimizing Harm and Maximizing Benefit	Beneficence is not met, no matter how minimal the risks, when there is no maximized benefit to the tribe or its participants. This in turn can lead to an injustice if the benefits gained by that research are denied to the tribe and/or its citizens. Cherokee Nation, (2019a) The Legislature also has a fundamental responsibility to protect and preserve the culture of the Nation and to ensure that the IRB permitted activities are conducted in a way that does no harm to the culture of the Nation. Ho-Chunk Nation, (2005)	Cultural harm to be prevented in research and maximization of benefits to be treated as core rather than incidental aspect of research
E2: For Justice	Both the researcher(s) and Tribe must bring equity to any research contract, agreement, or understanding. This includes finances, community knowledge, networks, personnel, and political or social power. Three Affiliated Tribes (n.d.) Community knowledge, networks, and personnel and political or social power are other forms of equity useful to a project. Each of these commodities has value and must be shared between the researchers and the Tribe if a proper agreement is to be formulated. Turtle Mountain Band of Chippewa Indians, (2014)	Unequal relations in Indigenous research to be acknowledged and joint efforts to be made by researchers and tribes to address inequities through sharing of power, people, knowledge, and resources
E3: For Future Use	At a minimum, the following information shall be provided by an applicant researcher... whether secondary use of any retained specimens is contemplated; informed consent regarding saved specimens and future uses... Ho-Chunk Nation, (2005) What control will the Community or Medical and Health Care Research participants have over the current and future use of the data, and how will the control be exercised?... What control will the Community have over the current and future use of the human biological material, and how will the control be exercised? (9.107) Gila River Indian Community, (2009)	Disclosure, consent, and control required with respect to secondary uses of research materials and data

Indigenous innovation and self-determination by focusing on people and purpose-oriented standards to be used alongside mainstream data guidelines (Carroll et al., 2020).

Tribal Research Governance as Expectations

The CARE Principles are in the early stages of implementation, with some entities leading the way by collaborating with the Global Indigenous Data Alliance (GIDA) to operationalize the principles within repositories, national ethics frameworks, and United Nations open science guidance (Australian Institute of Aboriginal and Torres Strait Islander Studies, 2020; Carroll et al., 2021; United Nations Educational Scientific and Cultural Organization, 2021; Welch et al., 2021). Large international genomics consortia are already implementing the FAIR principles, but to truly engage and demonstrate respect for marginalized, impacted, excluded, underserved populations, the CARE Principles must be integrated across institutional policies and practices (Wilkinson et al., 2016; Carroll et al., 2021). We draw on federal- and state-recognized Native nations' research regulations (**Table 2**) to illustrate how these official documents' assertions of IDSov set tribal expectations for enacting the CARE Principles. Tribal expectations include alignment with tribal priorities, recognizing the locus of control for tribal data, supporting respectful relationships, and addressing inequities in research.

DISCUSSION AND RECOMMENDATIONS

Below we make recommendations for tribes, other institutions, and ethical practices that leverage Native nations' codes as standards for researchers and data stewards as they implement the CARE Principles.

Tribal Law and Policy

Native nations are increasingly using tribal codes to set standards and expectations, exerting their jurisdiction over data, interests, places, and issues both on and off reservations (National Congress of American Indians, 2019a; Hiraldo et al., 2020). Here we share some of the ways that tribes address some of the more complex issues of tribal research oversight, including jurisdiction off tribal lands and protection of individual and collective interests, to spur Native nations to create and strengthen codes as guides to use of the CARE Principles with their peoples, lands, knowledges, and resources.

The fact that most tribal citizens reside off tribal lands (Norris et al., 2012), but may participate in research, raises unique challenges to the exercise of tribal sovereignty in research. Tribes have sought to address this governance challenge by extending the application of their research codes beyond tribal lands in two situations: (1) use of materials to which tribes have a legal claim and (2) participation of tribal citizens. Some tribes extend the protection of their citizens and interests beyond their territories by linking the exercise of their sovereignty to the physical location of research materials to which they have a claim (Colorado River Indian Tribes, 2009; Gila River Indian Community, 2009; Sisseton-Wahpeton Oyate Tribe, n.d.; no date, henceforth n.d.; United Houma Nation Institutional Review Board Ordinance, n.d.; White Earth Nation, 2018; Tribal Collaboration Working Group, 2018). Other tribes address research governance challenges beyond their territories by linking the exercise of sovereignty to participation of their citizens in research, particularly in studies that implicate aspects of their tribal citizenship and affiliation in some way (Navajo Nation, 2002; Confederated Tribes of Siletz Indians, 2005; Ho-Chunk Nation, 2005; Pascua Yaqui Tribe, 2008).

Some tribal claims of ownership over specimens and data are made in the context of broader statements about tribal sovereignty. For example, the Three Affiliated Tribes (n.d.) includes a general principle of prior rights that recognizes, among other rights, "proprietary rights and interests over… all knowledge and intellectual property" associated with their resources. Similarly, the United Houma (n.d.) Institutional Review Board Ordinance codifies the rights of the Tribe, "as a self-governed and self-determined people," to "all data and information generated and produced by… research" conducted in the community. Other codes couch the tribe's claim to ownership of specimens and data in narrower terms (Pascua Yaqui Tribe, 2008; Confederated Tribes of Siletz Indians, 2005; Sisseton-Wahpeton Oyate Tribe, n.d.), while others stress the need for researchers to respect those claims (Mohawk Nation of Akwesasne, 1996; Cherokee Nation, 2019b).

Some codes protect not only tribal (i.e., collective) but also individual citizens' claims to ownership and control of specimens and data (Tohono O'odham Nation, 2013; Colorado River Indian Tribes, 2009). Tribes have adopted intellectual property provisions in their codes to support individual and collective claims of ownership in specimens and data (Mohawk Nation of Akwesasne, 1996; Navajo Nation, 2002; Ho-Chunk Nation, 2005; Colorado River Indian Tribes, 2009; Turtle Mountain Band of Chippewa Indians, 2014).

Issues in research agreements pertaining to data reflect broad tribal concerns about specimens. Additional points include the need to: describe specific means of preserving confidentiality of individual and tribal data, including Assurances of Confidentiality (Mohawk Nation of Akwesasne, 1996; Ho-Chunk Nation, 2005; Gila River Indian Community, 2009); provide data disposal plans (Cherokee Nation, 2019b); and detail conditions that would allow researchers to breach their duty of confidentiality under signed agreements (Mohawk Nation of Akwesasne, 1996; Ho-Chunk Nation, 2005).

International, Federal, and Institutional Guidelines

As institutions increasingly operationalize the CARE Principles in policy and practice, understanding how high-level principles link to tribal expectations becomes paramount. While research institutions, researchers, and funding agencies must follow appropriate federal, state, and local laws, they must also follow proper engagement and consultation procedures with tribal nations to uphold tribal law and policy pertaining to research, data, and specimens. Tribal laws and processes need to be part of robust planning and policy for research institutions and programs to implement the CARE Principles. Importantly, each Native nations' written standards apply to research relationships with that nation only. The written standards must be balanced with ongoing community

relationships to give more depth and definitions to community expectations and needs. Additionally, when no laws exist, it is the responsibility of research institutions, researchers, and funding agencies to engage in a process with participating tribal nations to obtain approvals and guidance for research and data oversight (Tribal Collaboration Working Group, 2018; National Congress of American Indians, 2019a). Finally, examining the commonalities across Native nations provides insight into broad and common ethical expectations.

Evolving Ethical Practices

The CARE Principles, especially as indicated by tribal research codes, delineate standards for research practice. Training for researchers to understand tribal sovereignty, tribal codes, and review processes is necessary to provide the knowledge and tools to meet tribal ethical expectations. Supporting the CARE Principles requires an approach to biomedical research and policy that supports tribal ethics requirements, regardless if they have been codified as law.

Institutions, researchers, tribes, and Indigenous communities will benefit from careful attention to the CARE Principles to enhance trust and build meaningful relationships to ensure high quality translational biomedical science that emerges as tangible benefits for tribes and rural and urban Indigenous communities.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SC and NG conceptualized, drafted, and finalized the manuscript. IG, VH, MH, RP, and DSR contributed to drafting and editing.

FUNDING

The Morris K. Udall and Stewart L. Udall Foundation supported SC, IG, and RP.

ACKNOWLEDGMENTS

The authors acknowledge Indigenous Peoples worldwide, and the tribes in the US that are re-imagining and re-membering laws, policies, and practices to assert their right to control and enact their responsibility to protect data about their peoples, lands, and resources. Special appreciation to the Native nations whose codes and policies are cited in this document for making this information publicly available so others are able to learn from it. Some concepts in this essay emerged as comments in response to the Request for Information (RFI) soliciting additional input to the All of Us Research Program 2019 Tribal Consultation (Notice Number: NOT-PM-19-004). SC is also grateful for the thoughtful comments of Professor Jane Anderson, and for the invitation to present this research to the "Innovation Policy Colloquium" at New York University School of Law. The authors are grateful to Danella Hall, Candace Yazzie, and Rogena Sterling for related work that spurred ideas and concepts.

REFERENCES

Around Him, D., Andalcio Aguilar, T., Frederick, A., Larsen, H., Seiber, M., and Angal, J. (2019). Tribal IRBs: A Framework for Understanding Research Oversight in American Indian and Alaska Native Communities. *Am. Indian Alsk. Native Ment. Health Res.* 26 (2), 71–95. doi:10.5820/aian.2602.2019.71

Australian Institute of Aboriginal and Torres Strait Islander Studies (2020). *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research*. Canberra, Australia: AIATSIS. Available at: https://aiatsis.gov.au/sites/default/files/2020-10/aiatsis-code-ethics.pdf (Accessed January 19, 2022).

Boyer, B. B., Mohatt, G. V., Pasker, R. L., Drew, E. M., and McGlone, K. K. (2007). Sharing Results from Complex Disease Genetics Studies: a Community Based Participatory Research Approach. *Int. J. Circumpolar Health* 66 (1), 19–30. doi:10.3402/ijch.v66i1.18221

Carroll, S. R., Rodriguez-Lonebear, D., and Martinez, A. (2019). Indigenous Data Governance: Strategies from United States Native Nations. *Data Sci. J.* 18 (31), 1–15. doi:10.5334/dsj-2019-031

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materrechera, S., et al. (2020). The CARE Principles for Indigenous Data Governance. *Data Sci. J.* 19 (43), 1–12. doi:10.5334/dsj-2020-043

Carroll, S. R., Herczog, E., Hudson, M., Russell, K., and Stall, S. (2021). Operationalizing the CARE and FAIR Principles for Indigenous Data Futures. *Sci. Data* 8 (1), 1–6. doi:10.1038/s41597-021-00892-0

Chadwick, J. Q., Copeland, K. C., Branam, D. E., Erb-Alvarez, J. A., Khan, S. I., Peercy, M. T., et al. (2019). Genomic Research and American Indian Tribal Communities in Oklahoma: Learning from Past Research Misconduct and Building Future Trusting Partnerships. *Am. J. Epidemiol.* 188 (7), 1206–1212. doi:10.1093/aje/kwz062

Cherokee Nation (2019a). Guidance on Study Closure. Information and Guidance Documents. Available at: https://irb.cherokee.org/media/jezhjftz/guidance-on-study-closure.pdf (Accessed January 19, 2022).

Cherokee Nation (2019b). Tribal Research. Information and Guidance Documents. Available at: https://irb.cherokee.org/media/zaulx5iz/tribal-reserach.pdf (Accessed January 19, 2022).

Colorado River Indian Tribes (2009). Human and Cultural Research Code. Ordinance No. 09-04. Available at: https://www.crit-nsn.gov/crit_contents/ordinances/Human-and-Cultural-Research-Code.pdf (Accessed January 19, 2022).

Confederated Tribes of Siletz Indians (2005). Research Ordinance. Ordinance No. 9.100. Available at: https://www.ctsi.nsn.us/wp-content/uploads/2020/12/Research-Ordinance-09-16-2005.pdf (Accessed January 19, 2022).

Department of the Interior (2021). Indian Entities Recognized and Eligible to Receive Services from the United States Bureau of Indian Affairs. Federal Register 86, No. 67. 18552. Available at: https://www.govinfo.gov/content/pkg/FR-2021-04-09/pdf/2021-06723.pdf (Accessed January 19, 2022).

Dirks, L. G., Shaw, J. L., Hiratsuka, V. Y., Beans, J. A., Kelly, J. J., and Dillard, D. A. (2019). Perspectives on Communication and Engagement with Regard to Collecting Biospecimens and Family Health Histories for Cancer Research in a Rural Alaska Native Community. *J. Community Genet.* 10 (3), 435–446. doi:10.1007/s12687-019-00408-9

Duthu, N. B. (2008). *American Indians and the Law*. London: Penguin Publishing Group.

Garrison, N. A., Hudson, M., Ballantyne, L. L., Garba, I., Martinez, A., Taualii, M., et al. (2019a). Genomic Research through an Indigenous Lens: Understanding the Expectations. *Annu. Rev. Genom. Hum. Genet.* 20 (1), 495–517. doi:10.1146/annurev-genom-083118-015434

Garrison, N. A., Barton, K. S., Porter, K. M., Mai, T., Burke, W., and Carroll, S. R. (2019b). Access and Management: Indigenous Perspectives on Genomic Data Sharing. *Ethn. Dis.* 29 (Suppl. 3), 659–668. doi:10.18865/ed.29.S3.659

Gila River Indian Community (2009). The Medical and Health Care Research Ordinance (Ord GR-05-09), Title 17, Chapter 9. Gila River Indian Community Code. Available at: https://naair.arizona.edu/sites/default/files/gila_river_indian_ord_gr-05-09_title_17_chapter_9_0.pdf (Accessed January 19, 2022).

Haozous, E. A., Lee, J., and Soto, C. (2021). Urban American Indian and Alaska Native Data Sovereignty: Ethical Issues. *Am. Indian Alsk. Native Ment. Health Res.* 28 (2), 77–97. doi:10.5820/aian.2802.2021.77

Haring, R. C., Henry, W. A., Hudson, M., Rodriguez, E. M., and Taualii, M. (2018). Views on Clinical Trial Recruitment, Biospecimen Collection, and Cancer Research: Population Science from Landscapes of the Haudenosaunee (People of the Longhouse). *J. Canc Educ.* 33 (1), 44–51. doi:10.1007/s13187-016-1067-5

Hiratsuka, V. Y., Hahn, M. J., Woodbury, R. B., Hull, S. C., Wilson, D. R., Bonham, V. L., et al. (2020). Alaska Native Genomic Research: Perspectives from Alaska Native Leaders, Federal Staff, and Biomedical Researchers. *Genet. Med.* 22 (12), 1935–1943. doi:10.1038/s41436-020-0926-y

Ho-Chunk Nation (2005). Tribal Research Code, Ho-Chunk Nation Code, 3 HCC §3. Available at: https://tribalinformationexchange.org/wp-content/uploads/2019/03/Ho-Chunk-Nation-Tribal-Research-Code.pdf (Accessed January 19, 2022).

Hudson, M., Garrison, N. A., Sterling, R., Caron, N. R., Fox, K., Yracheta, J., et al. (2020). Rights, Interests and Expectations: Indigenous Perspectives on Unrestricted Access to Genomic Data. *Nat. Rev. Genet.* 21, 377–384. doi:10.1038/s41576-020-0228-x

James, R., Tsosie, R., Sahota, P., Parker, M., Dillard, D., Sylvester, I., et al. (2014). Exploring Pathways to Trust: a Tribal Perspective on Data Sharing. *Genet. Med.* 16 (11), 820–826. doi:10.1038/gim.2014.47

Karuk Tribe (2015). Protocol with Agreement for Intellectual Property Rights of the Karuk Tribe: Research, Publication and Recordings. Publication and Recordings. Available at: https://www.karuk.us/images/docs/forms/Protocol_with_Agreement_for_Intellectual_Property_Rights_of_the_Karuk_Tribe.pdf (Accessed January 19, 2022).

Kukutai, T., and Taylor, J. (Editors) (2016a). *Indigenous Data Sovereignty: Toward an Agenda*. Canberra, Australia: Australian National University Press.

Kukutai, T., and Taylor, J. (2016b). "Data Sovereignty for Indigenous Peoples: Current Practice and Future Needs," in *Indigenous Data Sovereignty: Toward an Agenda*. Editors T. Kukutai and J. Taylor (Canberra, Australia: Australian National University Press), 1–24.

Lovett, R., Lee, V., Kukutai, T., Rainie, S. C., and Walker, J. (2019). "Good Data Practices for Indigenous Data Sovereignty," in *Good Data*. Editors A. Daly, K. Devitt, and M. Mann (Amsterdam: Institute of Network Cultures).

Maiam nayri Wingara (2018). Indigenous Data Sovereignty Communique. Canberra, Australia: Maiam Nayri Wingaara. Available at https://static1.squarespace.com/static/5b3043afb40b9d20411f3512/t/5b6c0f9a0e2e725e9cabf4a6/1533808545167/Communique%2B-%2BIndigenous%2BData%2BSovereignty%2BSummit.pdf (Accessed January 19, 2022).

Mohawk Nation of Akwesasne (1996). Protocol for Review of Environmental and Scientific Research Proposals, Akwesasne Task Force on the Environment and Research Advisory Committee. Available at: https://nnigovernance.arizona.edu/protocol-review-environmental-and-scientific-research-proposals (Accessed January 19, 2022).

National Conference of State Legislatures (2019). Federal and State Recognized Tribes. Available at https://www.ncsl.org/research/state-tribal-institute/list-of-federal-and-state-recognized-tribes.aspx (Accessed January 19, 2022).

National Congress of American Indians (2019a). *Re: All of Us Research Program Tribal Consultation*. Washington, DC: National Congress of American Indians. Available at: http://www.ncai.org/policy-research-center/research-data/recommendations/NCAI_Letter_NIH_All_of_Us_Research_Program_Consultation_11_26_2019_FINAL_signed.pdf (Accessed January 19, 2022).

National Congress of American Indians (2019b). *Resolution ABQ-19-061 - Calling upon the National Institutes of Health to Consult with Tribal Nations and Establish Policies and Guidance for Tribal Oversight of Data on Tribal Citizens Enrolled in the* All Of Us *Research Program*. Albuquerque: National Congress of American Indians. Available at: https://www.ncai.org/ABQ-19-061.pdf (Accessed January 19, 2022).

National Congress of American Indians (2021). *Consultation Support Center*. Washington, DC: National Congress of American Indians. Available at: https://www.ncai.org/resources/consultation-support (Accessed January 19, 2022).

National Institutes of Health (2014). Genomic Data Sharing Policy. *Fed. Regist.* 79, 16751345–16751354. Available at: https://www.govinfo.gov/content/pkg/FR-2014-08-28/pdf/2014-20385.pdf (Accessed January 19, 2022).

Navajo Nation (2002). Human Research Code. Health and Welfare. 13 NNC §§3251 through 3271. April 16, 2002. Available at: https://www.nnhrrb.navajo-nsn.gov/pdf/NavNatHumResCode.pdf (Accessed January 19, 2022).

Norris, T., Vines, P. L., and Hoeffel, E. M. (2012). *The American Indian and Alaska Native Population: 2010*. Washington, D.C: Census Bureau. Available at: https://www.census.gov/prod/cen2010/briefs/c2010br-10.pdf (Accessed January 19, 2022).

Office of Human Research Protections (2017). Federal Policy for the Protection of Human Subjects. Final Rule. *Fed. Regist.* 82, 7149–7274. Available at: https://www.gpo.gov/fdsys/pkg/FR-2017-01-19/pdf/2017-01058.pdf (Accessed January 19, 2022).

Pascua Yaqui Tribe (2008). Research Protection. 8 PYTC §§7-1-10 through 7-1-250. May 14, 2008. Available at: https://www.pascuayaqui-nsn.gov/tribal-code/ch-7-1-research-protection/ (Accessed January 19, 2022).

Rainie, S. C., Rodriguez-Lonebear, D., and Martinez, M. (2017). *Policy Brief (Version 2): Data Governance for Native Nation Rebuilding*. Tucson: Native Nations Institute, University of Arizona. Available at: https://climas.arizona.edu/sites/default/files/Policy_Brief_Data_Governance_for_Native_Nation_Rebuilding_Version_2.pdf (Accessed January 19, 2022).

Rainie, S. R., Kukutai, T., Walter, M., Figueroa-Rodriguez, O. L., Walker, J., and Axelsson, P. (2019). "Issues in Open Data: Indigenous Data Sovereignty," in *The State of Open Data: Histories and Horizons*. Editors T. Davies, S. Walker, M. Rubinstein, and F. Perini (Cape Town and Ottawa: African Minds and International Development Research Centre), 300–319.

Research Data Alliance International Indigenous Data Sovereignty Interest Group (2019). "CARE Principles for Indigenous Data Governance," in *The Global Indigenous Data Alliance*. Editors S. R. Carroll and M. Hudson. Available at: https://www.gida-global.org/ (Accessed February 11, 2022).

Shaw, J. L., Robinson, R., Starks, H., Burke, W., and Dillard, D. A. (2013). Risk, Reward, and the Double-Edged Sword: Perspectives on Pharmacogenetic Research and Clinical Testing Among Alaska Native People. *Am. J. Public Health* 103 (12), 2220–2225. doi:10.2105/AJPH.2013.301596

Sisseton-Wahpeton Oyate Tribe (n.d.). Research Code. Sisseton-Wahpeton Oyate Codes of Law §§77-01-01 through 77-07-06. Available at: https://www.swo-nsn.gov/wp-content/uploads/Chapter-77-ResearchCode-FINAL-V15.pdf (Accessed January 19, 2022).

Smith, L. T. (2012). *Decolonizing Methodologies: Research and Indigenous Peoples*. Second edition. London: Zed Books.

Taitingfong, R., Bloss, C. S., Triplett, C., Cakici, J., Garrison, N., Cole, S., et al. (2020). A Systematic Literature Review of Native American and Pacific Islanders' Perspectives on Health Data Privacy in the United States. *J. Am. Med. Inform. Assoc.* 27 (12), 1987–1998. doi:10.1093/jamia/ocaa235

TallBear, K. (2013). *Native American DNA: Tribal Belonging and the False Promise of Genetic Science*. Minneapolis: University of Minnesota Press.

Three Affiliated Tribes: Mandan, Hidatsa, and Arikara Nation (n.d.). Research Code.

Tohono O'odham Nation (2013). Research Code. 17 Tohono O'odham Code Chapter 8. Available at: https://naair.arizona.edu/sites/default/files/tohono_oodham_nation_title17ch8_0.pdf (Accessed January 19, 2022).

Tribal Collaboration Working Group (2018). *All of Us Research Program Advisory Panel*. Considerations for Meaningful Collaboration with Tribal Populations. Available at: https://allofus.nih.gov/sites/default/files/tribal_collab_work_group_rept.pdf (Accessed January 19, 2022).

Trinidad, S. B., Ludman, E. J., Hopkins, S., James, R. D., Hoeft, T. J., Kinegak, A., et al. (2015). Community Dissemination and Genetic Research: Moving beyond Results Reporting. *Am. J. Med. Genet.* 167 (7), 1542–1550. doi:10.1002/ajmg.a.37028

Tsosie, R. (2019). Tribal Data Governance and Informational Privacy: Constructing 'Indigenous Data Sovereignty. *Mont. L. Rev.* 80 (2).

Turtle Mountain Band of Chippewa Indians (2014). Research Protection Act. Available at: https://nebula.wsimg.com/792a7ad79f2b2b8886d214cd7373ab8c?AccessKeyId=FF0318FE16C15B6C1F97&disposition=0&alloworigin=1 (Accessed January 19, 2022).

United Houma Nation (n.d.). Institutional Review Board Ordinance.

United Nations Educational Scientific and Cultural Organization (UNESCO) (2021). *UNESCO Recommendation on Open Science*. Paris: UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en (Accessed January 19, 2022).

Walter, M., Kukutai, T., Carroll, S. R., and Rodriguez-Lonebear, D. (Editors) (2021). *Indigenous Data Sovereignty and Policy*. New York: Routledge.

Welch, E., Louafi, S., Carroll, S. R., Hudson, M., IJsselmuiden, C., Kane, N., et al. (2021). *Post COVID-19 Implications on Genetic Diversity and Genomics*

Research & Innovation: A Call for Governance and Research Capacity. Rome: Food and Agricultural Organization of the United Nations. Available at: http://www.fao.org/3/cb5573en/cb5573en.pdf (Accessed January 19, 2022).

White Earth Nation (2018). Restructured/Codified White Earth Nation Research Review Board Policy/Procedure.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* 3, 160018. doi:10.1038/sdata.2016.18

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

PERSPECTIVE
published: 26 April 2022
doi: 10.3389/fgene.2022.795992

Precision Medicine Needs to Think Outside the Box

Daphne O. Martschenko[1]* *and Jennifer L. Young*[1,2]

[1]Center for Biomedical Ethics, School of Medicine, Stanford University, Stanford, CA, United States, [2]Center for Genetic Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Precision medicine offers a precious opportunity to change clinical practice and disrupt medicine's reliance on crude racial, ethnic, or ancestral categories by focusing on an individual's unique genetic, environmental, and lifestyle characteristics. However, precision medicine and the genomic studies that are its cornerstone have thus far failed to account for human diversity. This failure is made clearer when looking at individuals who encapsulate a mosaic of different genetic ancestries and do not fit neatly into existing population labels. This piece argues that precision medicine continues to rely on the same forms of crude categorization it seeks to unsettle. Until the scientific community creates inclusive solutions for individuals who fall outside or between our existing population labels, precision medicine will continue to fall short in its aims.

Keywords: precision medicine, GWAS, genome-wide association study, multiracial, admixed, equity, mixed ancestry

OPEN ACCESS

Edited by:
Arslan A. Zaidi,
University of Pennsylvania,
United States

Reviewed by:
Diddahally Govindaraju,
Albert Einstein College of Medicine,
United States
Tina Lasisi,
The Pennsylvania State University
(PSU), United States

***Correspondence:**
Daphne O. Martschenko
daphnem@stanford.edu

Specialty section:
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

Received: *15 October 2021*
Accepted: *07 April 2022*
Published: *26 April 2022*

Citation:
Martschenko DO and Young JL (2022) Precision Medicine Needs to Think Outside the Box. Front. Genet. 13:795992. doi: 10.3389/fgene.2022.795992

INTRODUCTION

An increasing number of individuals are defying the crude systems of racial, ethnic, and ancestral categorization used in medicine and society. For instance, over the past decade, the number of Americans who self-identify as multiracial has more than doubled (https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html) and increased globalization and population migration have resulted in greater genetic admixture—defined as the recent combination of two or more genetic ancestries. (Korunes and Goldberg, 2021). For those of us who do not fit neatly into existing racial, ethnic, or ancestral population labels, the problematic practice of categorizing people into discrete groups can be especially exclusionary. Precision medicine is one area in which such individuals are being left behind. In this Commentary, we argue that realizing the aims of precision medicine requires the medical genomics community to comprehensively study and analyze data from those who cannot be classified into existing population labels.

Precision Medicine

Precision medicine examines how an individual's unique genetic, environmental, and lifestyle characteristics come together to inform health. Instead of one-size-fits-all approaches to medical decisions, interventions, and treatments, precision medicine focuses on customization to the individual. Central to enabling such customization is medical genomics research–a heavily funded research priority for precision medicine (https://www.genome.gov/news/news-release/NHGRI-awards-73million-to-continue-building-Clinical-Genome-Resource-ClinGen). Researchers in medical genomics use genome-wide association studies (GWAS) to identify fine-grained differences in the DNA sequences of related and unrelated individuals. Aggregating the small effects of thousands of genetic variants identified through GWAS, polygenic scores (PGS) are used to estimate a person's likelihood of exhibiting a particular phenotype (e.g., cardiovascular disease).

Recent efforts in precision medicine have focused on how PGS might be used in combination with environmental risk scores (ERS) to screen individuals for diseases such as cancer.

The advent of precision medicine offers a precious opportunity to move beyond mutually exclusive categories such as race (Raut et al., 2021; Bonham et al., 2016) and account for human genetic diversity among individuals within and between populations. (Lewontin et al., 1972). GWAS and PGS could enable clinicians to make more accurate diagnoses and tailor treatments using an individual's genome instead of self-identified or inferred race or ethnicity. However, while precision medicine carries the promise of improving clinical care, preventing and treating disease, and rejecting the use of race-based corrections in medicine, (Ashley, 2015; Cerdeña et al., 2020), it has thus far failed to deliver. Such failures are made clearer when examining how medical genomics handles admixed individuals who encapsulate more than one genetic ancestry.

The Limitations of Precision Medicine for Admixed Individuals

Despite the rapidly decreasing costs associated with conducting GWAS, the overwhelming majority of genomic studies use samples from European genetic ancestries (https://gwasdiversitymonitor.com/); this restricts the potential benefits of genomics research on health to a narrow subset of the global population while also introducing sampling bias. (Popejoy and Fullerton, 2016). The challenges of population stratification, coupled with Euro-centric biases in genomic databases, mean that PGS derived from GWAS have systematically lower predictive performance when applied to understudied populations. As a result, the disease risk of non-European populations, including admixed populations, are either under- or over-estimated using existing PGS. (Martin et al., 2017). Any benefits afforded by PGS are less likely to accrue among people of non-European ancestry and more likely to exacerbate health disparities in disease treatment. (Martin et al., 2019).

In an effort to increase and diversify the sampling of participants, we must build databases that better reflect the global population, and widen the applicability of precision medicine research. Initiatives such as the NIH-funded *All of Us* Research Program are emerging (https://allofus.nih.gov/) to respond to this unmet need. However, initiatives such as these will never realize the full benefits of precision medicine unless explicit attention is devoted to finding ways to study admixed individuals in medical genomics research; this includes both existing admixed populations (e.g., Hispanic or Latin American) and recently admixed individuals who fall outside of already-defined admixed population categories.

Although genetics researchers are beginning to conduct studies with samples from diverse populations that encapsulate more than one genetic ancestry (e.g., self-identified African American or Hispanic/Latin American), (Wojcik et al., 2019; Gopalan et al., 2021), the vast majority of studies continue to deprioritize and discard admixed samples, citing inadequate sample sizes and technical complexities. (Peterson et al., 2019; Ben- Eghan et al., 2020).

These issues are further exacerbated for recently admixed individuals. First and second-generation admixed individuals are often grouped into monolithic categories such as 'Other admixed ancestry,' (Morales et al., 2018) 'Other and other admixed,' or 'Multiple' (https://www.ebi.ac.uk/gwas/docs/ancestry-data). Aggregating individuals into these categories may help to increase statistical power, but it denies researchers opportunities to examine relationships between the unique sociocultural factors and genetic characteristics that come together to shape an individual's health and well-being.

Current genomic methods are especially insufficient for analyzing data from first and second generation admixed individuals. Continental ancestry categories (e.g., European, African) are the most common type of group label in genomics research. (Panofsky and Bliss, 2017; Lewis et al., 2021). The overreliance on continental ancestry categories not only encourages dangerous slippage between genetic ancestry and race, (Panofsky and Bliss, 2017), it disincentivizes researchers from finding ways to include those who fall outside a broad continental grouping. For instance, an individual who is a recent combination of Greater Middle Eastern genetic ancestry and South East Asian genetic ancestry is likely to be categorized as 'Other and other admixed' and will be discarded from genomic analyses because they cannot be assigned to a distinct regional population grouping.

The current limitations of medical genomics raise important scientific and ethical considerations regarding missed scientific opportunities, underrepresentation in research, and participants' efforts to contribute to science. It is ethically problematic to continue inequitable resource allocations that drive underrepresentation in genomic studies, (Fatumo et al., 2022), just as it is ethically problematic to recruit participants for research and then discard their contributions from analyses. The consequence of such practices for precision medicine is that many do not currently stand to benefit from research into pharmacogenetics or disease risk prediction and will continue to be left behind even as the field outwardly seeks to diversify biobanks.

Possible Solutions

To address these issues, precision medicine must first recognize, incorporate, and amplify the work of researchers who are already grappling with issues of diversity and equity in clinical and healthcare contexts in and outside of genetics. (Panofsky and Bliss, 2017; Lewis et al., 2021). This means expanding the range of voices given decision-making capacities and committing to an ethos of diversity in research and the workplace. (McFarling, 2021)' (Thomas et al., 2021) Researchers must also prioritize community-engaged efforts that focus on building dynamic two-way partnerships instead of transactional exchanges for which data collection is an endpoint. Implementing more inclusive approaches to how precision medicine is carried out will introduce new perspectives and ways of thinking that can help to improve our current methods of analysis in genomics to account for admixed individuals.

In support of improving health outcomes and enhancing disease prevention and treatment, precision medicine should also consider whether existing systems of classification,

methodological approaches, and research priorities are appropriate. We join others in cautioning against our default use of continental ancestral groupings in genetics. (Lewis et al., 2021). Although admixed individuals, who are considered a mixture of broad continental groups, may be used to compound population labeling, (Lewis et al., 2021), we believe that admixed individuals such as ourselves offer a chance to escape from it. The limited framework for attaining diversity in genomics have negative consequences for those of us who do not fit into a box. Therefore a critical and reflexive audit of how precision medicine research is conducted, who it benefits, and the changes required, calls for additional specific attention to those who cannot be classified using our current population schema. Studying rather than ignoring recently admixed populations is not only a scientific and ethical imperative, it will provide opportunities to develop novel methods and analytic techniques that resist continental ancestry groups and help realize the full potential of precision medicine for all. (Peterson et al., 2019).

Finally, precision medicine initiatives must prioritize investigations of the social context and the role of social and environmental factors including structural racism in shaping human health. If we want precision medicine to benefit all and not just some, the research enterprise needs to understand the systems-level factors that contribute to health disparities. (Newman, 2021). Individuals who defy the crude systems of racial, ethnic, and ancestral categorization used in medicine and society carry unique lived experiences that cannot be captured by genetics alone. These experiences are shaped by social contexts and hold potentially important health implications. Understanding the multitude of ways that individuals who do not fit into a box experience health is critical to offering genuinely customizable healthcare.

CONCLUSION

Precision medicine is failing those who do not fit neatly within our crude systems of categorization—whether they be racial, ethnic, *or* genetic ancestral. The limitations of precision medicine for recently admixed individuals who cannot be described using existing population labels illustrate this. Precision medicine will not dismantle our reliance on reductionist categorizations by using the very tools that require them. And, it will not improve health outcomes with biased genomic databases that leave out large swaths of the global population and distract from the social structures and systems that contribute to health. Until we recognize the limitations of our approach to precision medicine and seriously grapple with who it leaves out, we cannot rely on it to systematically improve how we prevent and treat disease.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

DM and JY jointly conceived paper topic and key arguments. Both authors worked equally on earlier iterations of the manuscript before DM took primary responsibility for incorporating feedback and finalizing the manuscript. DM is lead author.

FUNDING

This manuscript is supported by grant T32HG008953 (The Stanford Training Program in ELSI Research) and grant 19SFRN34830054 (The American Heart Association Arrhythmias Sudden Cardiac Death Strategically Focused Research Network).

ACKNOWLEDGMENTS

We would like to thank Emily Greenwald and Sam Trejo for their feedback on this manuscript.

REFERENCES

Ashley, E. A. (2015). The Precision Medicine Initiative. *JAMA* 313, 2119–2120. doi:10.1001/jama.2015.3595

Ben-Eghan, C., Sun, R., Hleap, J. S., Diaz-Papkovich, A., Munter, H. M., Grant, A. V., et al. (2020). Don't Ignore Genetic Data from Minority Populations. *Nature* 585, 184–186. doi:10.1038/d41586-020-02547-3

Bonham, V. L., Callier, S. L., and Royal, C. D. (2016). Will Precision Medicine Move Us beyond Race? *N. Engl. J. Med.* 374, 2003–2005. doi:10.1056/nejmp1511294

Cerdeña, J. P., Plaisime, M. V., and Tsai, J. (2020). From Race-Based to Race-Conscious Medicine: How Anti-racist Uprisings Call Us to Act. *The Lancet* 396, 1125–1128.

Fatumo, S., Chikowore, T., Choudhury, A., Ayub, M., Martin, A. R., and Kuchenbaecker, K. (2022). A Roadmap to Increase Diversity in Genomic Studies. *Nat. Med.* 28, 243–250. doi:10.1038/s41591-021-01672-4

Gopalan, S., Smith, S. P., Korunes, K., Hamid, I., Ramachandran, S., Goldberg, A., et al. (2021). Human Genetic Admixture through the Lens of Population Genomics. *ArXiv210912190 Q-Bio*. doi:10.48550/arXiv.2109.12190

Korunes, K. L., and Goldberg, A. (2021). Human Genetic Admixture. *PLOS Genet.* 17, e1009374. doi:10.1371/journal.pgen.1009374

Lewis, A. C. F., Molina, S. J., Appelbaum, P. S., Fuentes, A., Ghosh, N., Dauda, B., et al. (2021). Getting Genetic Ancestry Right for Science and Society. *ArXiv211005987 Q-Bio*. doi:10.48550/arXiv.2110.05987

Lewontin, R. C. (1972). "The Apportionment of Human Diversity," in *Evolutionary Biology*. Editors T. Dobzhansky, M. K. Hecht, and W. C. Steere (Springer US), 6, 381–398. doi:10.1007/978-1-4684-9063-3_14

Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., et al. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* 100, 635–649. doi:10.1016/j.ajhg.2017.03.004

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., and Daly, M. J. (2019). Clinical Use of Current Polygenic Risk Scores May Exacerbate

Health Disparities. *Nat. Genet.* 51, 584–591. doi:10.1038/s41588-019-0379-x

McFarling, U. L. (2021). *'Health Equity Tourists': How white Scholars Are Colonizing Research on Health Disparities*. Boston, MA: STAT. Available at: https://www.statnews.com/2021/09/23/health-equity-tourists-white-scholars-colonizing-health-disparities-research/.

Morales, J., Welter, D., Bowler, E. H., Cerezo, M., Harris, L. W., McMahon, A. C., et al. (2018). A Standardized Framework for Representation of Ancestry Data in Genomics Studies, with Application to the NHGRI-EBI GWAS Catalog. *Genome Biol.* 19, 21. doi:10.1186/s13059-018-1396-2

Newman, A. M. (2021). Moving beyond Mistrust: Centering Institutional Change by Decentering the white Analytical Lens. *Bioethics* 36, 267–273. doi:10.1111/bioe.12992

Panofsky, A., and Bliss, C. (2017). Ambiguity and Scientific Authority. *Am. Sociol. Rev.* 82, 59–87. doi:10.1177/0003122416685812

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., et al. (2019). Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* 179, 589–603. doi:10.1016/j.cell.2019.08.051

Popejoy, A. B., and Fullerton, S. M. (2016). Genomics Is Failing on Diversity. *Nature* 538, 161–164. doi:10.1038/538161a

Raut, J. R., Schöttker, B., Holleczek, B., Guo, F., Bhardwaj, M., Miah, K., et al. (2021). A microRNA Panel Compared to Environmental and Polygenic Scores for Colorectal Cancer Risk Prediction. *Nat. Commun.* 12, 4811. doi:10.1038/s41467-021-25067-8

Thomas, S. P., Amini, K., Floyd, K. J., Willard, R., Wossenseged, F., Keller, M., et al. (2021). Cultivating Diversity as an Ethos with an Anti-racism Approach in the Scientific enterprise. *Hum. Genet. Genomics Adv.* 2, 100052. doi:10.1016/j.xhgg.2021.100052

Wojcik, G. L., Graff, M., Nishimura, K. K., Tao, R., Haessler, J., Gignoux, C. R., et al. (2019). Genetic Analyses of Diverse Populations Improves Discovery for Complex Traits. *Nature* 570, 514–518. doi:10.1038/s41586-019-1310-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

PERSPECTIVE
published: 26 April 2022
doi: 10.3389/fgene.2022.880170

Recommendations for Sustainable Ancient DNA Research in the Global South: Voices From a New Generation of Paleogenomicists

Maria C. Ávila-Arcos[1]†, Constanza de la Fuente Castro[2]*†, Maria A. Nieves-Colón[3]*† and Maanasa Raghavan[2]*†*

[1]International Laboratory for Human Genome Research, Universidad Nacional Autónoma de México, Querétaro, Mexico, [2]Department of Human Genetics, University of Chicago, Chicago, IL, United States, [3]Department of Anthropology, University of Minnesota Twin Cities, Minneapolis, MN, United States

OPEN ACCESS

Edited by:
Arslan A Zaidi, University of Pennsylvania, United States

Reviewed by:
Ripan Malhi, University of Illinois at Urbana-Champaign, United States

***Correspondence:**
Maria C. Ávila-Arcos
mavila@liigh.unam.mx
Constanza de la Fuente Castro
cdelafuente@uchicago.edu
Maria A. Nieves-Colón
mnievesc@umn.edu
Maanasa Raghavan
mraghavan@uchicago.edu

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

Received: *21 February 2022*
Accepted: *23 March 2022*
Published: *26 April 2022*

Citation:
Ávila-Arcos MC, de la Fuente Castro C, Nieves-Colón MA and Raghavan M (2022) Recommendations for Sustainable Ancient DNA Research in the Global South: Voices From a New Generation of Paleogenomicists. Front. Genet. 13:880170. doi: 10.3389/fgene.2022.880170

Paleogenomics - the study of ancient genomes - has made significant contributions, especially to our understanding of the evolutionary history of humans. This knowledge influx has been a direct result of the coupling of next-generation sequencing with improved methods for DNA recovery and analysis of ancient samples. The appeal of ancient DNA studies in the popular media coupled with the trend for such work to be published in "high impact" journals has driven the amassing of ancestral human remains from global collections, often with limited to no engagement or involvement of local researchers and communities. This practice in the paleogenomics literature has led to limited representation of researchers from the Global South at the research design and subsequent stages. Additionally, Indigenous and descendant communities are often alienated from popular and academic narratives that both involve and impact them, sometimes adversely. While some countries have safeguards against 'helicopter science', such as federally regulated measures to protect their biocultural heritage, there is variable oversight in others with regard to sampling and exportation of human remains for destructive research, and differing requirements for accountability or consultation with local researchers and communities. These disparities reveal stark contrasts and gaps in regional policies that lend themselves to persistent colonial practices. While essential critiques and conversations in this sphere are taking place, these are primarily guided through the lens of US-based heritage legislation such as the Native American Graves and Protection Act (NAGPRA). In this article, we aim to expand the scope of ongoing conversations by taking into account diverse regional contexts and challenges drawing from our own research experiences in the field of paleogenomics. We emphasize that true collaborations involve knowledge sharing, capacity building, mutual respect, and equitable participation, all of which take time and the implementation of sustainable research methods; amass-and-publish strategy is simply incompatible with this ethos.

Keywords: paleogenetics, sustainable research, regulation, global south, diversity, accountability

INTRODUCTION

The field of ancient DNA (aDNA) has grown from a marginal subject that faced early scepticism, to a highly attractive and expanding field that has produced significant knowledge on the evolutionary history of several species, especially our own (Orlando et al., 2021). Numerous methodological improvements, including next generation sequencing, have considerably increased the yield of authentic aDNA from degraded biological materials. This is advantageous because aDNA research requires the destruction of valuable tissues or materials, making it imperative to handle this limited resource ethically, with legitimate justification, and to secure the recovery of as much information as possible from small sample sizes.

A breakthrough in human aDNA research was the development of "capture" technology, characterizing a subset of approximately 1,240,000 genome-wide variant sites ('1240k' panel) that are primarily informative for population history inference. Notably, the capture approach was conceptually proposed in 2014 as a strategy to "democratize" the field by offering a cost-effective alternative to whole-genome/shotgun sequencing (Pickrell and Reich, 2014). However, for commercial reasons, this capture assay was initially not publicly accessible to all research groups, which ironically did the opposite of democratization by concentrating the power of this method to groups collaborating with its developers and early adopters.

The cost benefits of this method incited the few research groups with access to this technology and those with large budgetary resources to seek and stockpile ancient human samples. The flames of this "bone rush" (Fox and Hawks, 2019) were further fanned by the sensationalization of aDNA findings via "high impact" publications and popular media, raising a number of ethical concerns (Lewis-Kraus, 2019). Many of these sample collection campaigns were initially done without strong scrutiny by the scientific community. However, in recent years, critics have raised concerns about this "grab-and-go" approach, calling out its extractive nature and lack of engagement or meaningful involvement with local researchers and communities (e.g. (Bardill et al., 2018; Claw et al., 2018; Hudson et al., 2020; Tsosie et al., 2020; Wagner et al., 2020; Argüelles et al., 2022), more references in [1]). This practice is particularly conflictive when involving the collection and destruction of samples from nations in the "Global South"–a term often used to identify lower-income countries, many of which have been historically oppressed by colonialism (Dados and Connell, 2012)–by laboratories in the "Global North" (the complementary set of countries, many of which earned their higher wealth by colonization and exploitation of "Global South" nations).

Most criticisms of such unethical practices in paleogenomics research focus on United States contexts, the Native American Graves and Protection Act (NAGPRA) and Indigenous rights (Claw et al., 2018; Fox and Hawks, 2019; Wagner et al., 2020). Although a piece recommending 'global guidelines' for aDNA research was recently published, it was received with concern by some academics (Somel et al., 2021; Tsosie et al., 2021). As aDNA researchers from, and doing research in, Global South regions, we aim at expanding the discussion further by drawing from our own experiences to contribute to the ethical development of the paleogenomics field in Chile, India, Mexico, and Puerto Rico. We focus on four common challenges faced in our efforts to develop aDNA research programs anew as well as partner with existing programs in these places: 1) Cultural and heritage management regulations, 2) Local funding and infrastructure, 3) Local research and training capacity building, and 4) Consultation with Indigenous and descendant communities. We discuss these issues in the context of the aforementioned countries and outline recommendations from our experiences to address them, though admittedly not the only strategies for doing so.

In writing this piece, we acknowledge our positionality as researchers who may or may not share the same identities and histories with the communities with which we collaborate. We also recognize that our privileges as scientists impact our work and access to resources. The perspectives we discuss here represent our ongoing efforts as we learn, together with our community partners and trainees, how to build more sustainable, fair, and representative aDNA research programs globally.

Cultural and Heritage Management Regulations

Regulations surrounding destructive sampling and sample export for aDNA research projects vary widely and are enforced at different levels within governmental cultural or heritage institutions (Marquez-Grant and Fibiger, 2013). Regulatory bodies within these institutions are usually tasked with evaluating formal written requests to access collections and perform sampling and assure compliance with final reporting requirements. Besides institutions that enforce federal and/or state heritage regulations, in countries such as the United States and Canada, some Indigenous communities have their own regulations for genetic research studies (Claw et al., 2018, 2021; Begay et al., 2019; Wagner et al., 2020).

However, the reality in the Global South can be very different. Institutions tasked with regulating access to heritage or archaeological collections may lack or have unclear regulations and requirements, insufficient budgets or enforcement strategies to enact these regulations, or a combination thereof (e.g. (Abarca Labra et al., 2018)). This ambiguity can also affect local institutions or individuals (e.g. archaeologists leading excavations) who first receive research proposals for destructive analysis and where the sampling process itself may not be clearly defined nor accountability or follow-up measures outlined, leaving important decisions about sampling and research practices in the hands of a few.

Within Latin America, there is large variation in regulations for destructive sampling and few that are particularly dedicated to aDNA. For example, both Mexico and Chile have specific

[1]https://elsihub.org/index.php/collection/game-bones-power-ethics-and-emerging-technology-paleogenomics-research.

institutions dedicated to the research, conservation and protection of anthropological, archaeological, historical, and cultural heritage (National Institute of Anthropology and History (INAH) and National Monuments Council, respectively). INAH's Archaeology Council regulates access to archaeological samples for all destructive and non-destructive analyses following institutional guidelines (Instituto Nacional de Antropología e Historia, 2019). Although the council has clear requirements for ancient bio-molecular research, some of them are not ready to be implemented (e.g., lack of infrastructure to store aDNA extracts, libraries, or genetic data). In some countries from this region, cultural artifacts and human bodies recovered from archaeological contexts are subject to a heritage process where there is a centralized entity managing their fate. This process primarily involves legal protection against destruction and variable levels of regulations for research settings, particularly if samples are leaving the country. Albeit necessary, there are some troubling assumptions under this model that are rarely discussed by aDNA researchers, particularly in relation to regulatory and state recognition of Indigenous identities, which we discuss in following sections.

In India, the Archaeological Survey of India (ASI) serves as the premier national governmental institution that oversees the cultural heritage of the country, including regulating the export of materials abroad for research. Additionally, state government archaeological departments have the autonomy to carry out the role of heritage conservation and conduct archaeological excavations in their respective jurisdictions. These institutions are well suited to implement and oversee a best-practice regulatory framework for sampling of human bodies from archaeological contexts for aDNA research (Jamir, 2022b; Pappu, 2022; Rai, 2022; Taher, 2022).

Historically unequal power relations can also shape heritage management in ways that have consequences for aDNA research as seen in Puerto Rico, which as a U.S. territory without federally recognized Indigenous nations, is excluded from NAGPRA (Rodríguez López, 2009b; d'Alpoim Guedes et al., 2021). While artifacts and human bodies recovered from archaeological contexts are considered the patrimony of all Puerto Ricans under local legislation, such laws are superseded by federal regulations (Pagán-Jiménezand Rodriguez Ramos, 2008; Rodríguez López, 2009a;, 2009b; Llorens-Liboy and Núñez, 2011). Indeed, many archaeological remains excavated in Puerto Rico were exported to the US mainland soon after the American invasion (Pagán Jiménez, 2000); well before Puerto Ricans could vote for their own government or enact modern heritage legislation. Now curated in museums and collections abroad (DaRos and Colten, 2009; Françozo and Strecker, 2017), these remains can be legally sampled for aDNA research without passing through the permitting process required by Puerto Rican authorities or consulting with island stakeholder communities.

Recommendation: Support efforts to improve local cultural and heritage management regulations and involve these institutions in the research process. Researchers can aid communities, permitting agencies, and ethical and regulatory boards seeking to develop better frameworks and guidelines for destructive aDNA analyses by holding open discussions with board members about the process, risks and potential benefits of aDNA research, and by providing detailed reports and inventories of the remaining DNA products and data files after research is concluded. Furthermore, engaging with local museums or other heritage management institutions may facilitate contacts with communities for consultations prior to study start and assist with dissemination of research results to the general public, local museums, universities or schools.

Local Research and Training Capacity Building

Although building local capacities for research is an essential first step towards ensuring sustainability, equity and inclusion within paleogenomics, many Global South nations face significant challenges in setting up training programs and maintaining research facilities. These challenges include economic austerity measures that underfund public education and scientific investment, small or nonexistent research funding streams, lack of support for research capacity building, and limited access to, or structural disparities in, higher education (Reidpath and Allotey, 2019; Reidpath and Allotey, 2020; Viera et al., 2020; Carter and Hujo, 2021).

In Latin America and the Caribbean, for example, enrollments in higher education have increased over the last decade but other inequities persist, such as patterned access to higher education based on income and low availability of coursework or degree programs in science fields (Ferreyra et al., 2017). For example, despite being a US territory, Puerto Rico's universities are underfunded and about 42.9% of undergraduate students live below the poverty line (Nazario, 2014; Trines, 2018). As of this writing, there are limited opportunities for local undergraduate training in anthropology (Pagán Jiménez, 2000; Pagán-Jiménezand Rodriguez Ramos, 2008), and no formal graduate degree programs focused on biological anthropology, bioarchaeology, ancient DNA, or genomic science. While such programs can be pursued abroad, they often charge high tuition rates, making them unaffordable for many families, or limiting access just to high-income students. In Mexico, where a few degree programs in anthropology and genomics provide training in paleogenomics methods, severe budget cuts and government divestment threaten to reduce offerings and shutter educational institutions (Santos Cid, 2022). In India and Chile, degree programs in anthropology, archaeology, and genomics exist; however, there is currently only one functional aDNA lab in each country to train and conduct all local paleogenomics research.

Altogether, these factors make it difficult for local students to access and complete degree programs, reduce local job opportunities for scientific professionals, and accelerate brain-drain emigration patterns (Mishra, 2006; Weinberg, 2011; Docquier, 2014). In such contexts, local students may understandably see Global North countries as the only options for training, and local researchers may choose to export aDNA samples to these locations for processing and analyses. However, this creates a vicious cycle, as export means no local research is conducted, no training of students takes place, and local capacity for research remains underdeveloped.

TABLE 1 | Comparison of costs and delivery times for select consumables and reagents commonly used in aDNA research involving genome-wide next-generation sequencing. Cost estimates include tax, customs duties and shipping fees. Delivery times are the average number of days between ordering and arrival.

Country	Chile		India		Mexico		United States (UChicago)	
Reagent/Consumable	**Cost (USD)**	**Time (days)**	**Cost (USD)**	**Time (days)**	**Cost (USD)**	**Time (days)**	**Cost (USD)**	**Time (days)**
Filter tips (one set of P2/P10,P100/P200 and P1000)	$30	7	$75-$100	30–60	$29	14–60	$17	7
Sequencing (cost per Gb on the platform used most by local aDNA facility)	NextSeq550 2 × 75 (18 Gb); $100	30	HiSeq 4000 (120 Gb); $25	30–60	NextSeq550 2 × 75 (18 Gb); $82	21–30	NovaSeq SP 1 × 100 (45 Gb); $38	7–14
dsDNA library reagents; 1. NEBNext End Repair (E6050S); 2. NEBNext Quick Ligation E6056S; 3. Bst Polymerase Large Fragment M0275A	1. $172;2. $607; 3. $137	14–21	1. $100;2. $260; 3. $100	120–180	1. $196;2. $624; 3. $132	20–30	1. $74;2. $262;3. $59	7

Recommendation: Provide training opportunities to strengthen local research capacity. Researchers from Global North institutions can fund bilateral trainee exchanges with local institutions and provide wet lab and bioinformatics training in their labs. As shown by the success of the SING consortium workshops in the United states, Canada, Australia and Aotearoa (Malhi and Bader, 2015; Wade, 2018), short courses and training events can also become spaces to discuss multiple aspects of genomics research, including sampling strategies, data analyses or interpretations, and ethical considerations with local students and community members.

Local Funding and Infrastructure

Ancient DNA research requires sterile conditions for data generation and high data quality standards. This calls for an entirely separate facility dedicated to low copy number DNA processing that complies with strict criteria (Knapp et al., 2012; Fulton and Shapiro, 2019), and uses sterile molecular grade reagents and consumables (Champlot et al., 2010; Llamas et al., 2017). In the last 5–6 years, paleogenomics data generation has been overtaken by the '1240k' capture panel, costing nearly $250 per reaction without accounting for shipping and other export regulatory costs. Despite its early branding as "democratizing" paleogenomics (Pickrell and Reich, 2014), the capture panel remains out of reach for many laboratories in the Global South and may become yet another means of unequitable foreign collaborations (Argüelles et al., 2022).

Altogether, the cost of establishing, maintaining, and day-to-day running of a paleogenomics facility is not an easy feat, often requiring institutional and/or governmental commitment to infrastructure (e.g., lab space, equipment) and sustenance (e.g., hiring, reagents, consumables, maintenance). Not surprisingly, paleogenomics is not a high priority research avenue in most countries battling more pressing challenges, such as health crises or economic insecurity (Maher et al., 2012; Lebel and McLean, 2018; Liverpool, 2021). This often translates to little to no support for local researchers interested in developing this field in their countries, as reflected in the current distribution of global aDNA laboratories[2]. For countries in the Global South that have aDNA laboratories, upkeep is often difficult with issues ranging from power shortages to infrastructural and maintenance limitations such as limited space and scope for expansion, delayed and pricey access to equipment and parts, lack of expertise to diagnose and repair breakdown of proprietary equipment, and so on. Moreover, sourcing and ordering reagents that are easily obtained in the Global North is time-consuming and often several fold more expensive elsewhere (**Table 1**) (Ciocca and Delgado, 2017; Valenzuela-Toro and Viglino, 2021). When adding publication and conference fees, and the US dollar advantage, the cost of conducting research becomes

[2]https://www.google.com/maps/d/u/0/viewer?mid=1qwXOKV5uoQntgBsxQrxS01YHpbs&ll=52.19802207086798%2C-5.195632878906142&z=5 and https://isogg.org/wiki/List_of_forensic_and_ancient_DNA_laboratories.

disproportionately higher in the Global South vis-à-vis available budgets. Ultimately, these disparities in resource availability for executing scientific programs creates opportunities for 'grab-and-go' strategies that further research inequity and helicopter science practices (Adame, 2021; Haelewaters et al., 2021).

Recommendation: Foster equitable collaborations by supporting local research capacity and involving knowledgeable local collaborators and researchers as equal partners at all stages. Researchers based at Global North institutions can do this at different levels, from sharing equipment and reagents to local laboratories to formally collaborating with local researchers on projects and international grants. If collaborating in countries that already have aDNA facilities, additional ways to support capacity are to write research grants together with local collaborators, and budget for both reagents and consumables to be ordered to the local lab and travel to process samples jointly in the local laboratory and plans to contribute to training. The expertise and experience of local researchers and institutions can contribute nuanced insights and guide research goals to focus on locally relevant questions. Local collaborators are likely to have a better understanding of the regional history, current socio-political situation, regulations and, importantly, ethical implications of research for local communities. Partnerships with local collaborators and institutions can also aid in developing culturally responsive materials (Judd and McKinnon, 2021), in their language, for dissemination of research results to the public via news sources or other venues. This is important as genetics papers are often loaded with scientific jargon and assumptions that may be hard to interpret and open to misinterpretation by those far from the research (Harmon, 2018; Reich, 2018; Gannon, 2019; Wolinsky, 2019; Panofsky et al., 2021).

Consultation with Indigenous and Descendant Communities

In the United States, NAGPRA legally requires researchers to identify affiliated tribal nations for consultation about research with ancestral remains. Meanwhile, in many Global South countries, in addition to the lack of legal mandates for community consultation, there are other issues surrounding unclear regulations on heritage management (discussed above), Indigenous identities, and heterogeneity in the state-Indigenous dynamics across and within regions that should be considered while enacting nuanced, community-sensitive consultation and engagement strategies.

Ethnic identity and belonging are fluid sociocultural constructs with definitions that vary over time and between populations. In some Indigenous communities, these constructs can be tied to connections with land or ancestors, and with spiritual, cultural, religious and linguistic traits, while others may invoke biological (phenotypic) features. Insights from genetic ancestry studies, if not framed sensitively and acknowledging existing identity structures, could impose upon the process of identity construction for both individuals and communities (Egorova, 2009; Gibbon et al., 2011; TallBear, 2013; Wade et al., 2015; Benn Torres and Torres Colón, 2020; Crellin and Harris, 2020). Genetic insights can sometimes conflict with community and individual beliefs or reproduce nationalistic or essentialized notions of identity that suppress the existence of Indigenous peoples.

In some Latin American countries, including Mexico and Chile, the institutional de-indigenization processes put in practice by the state emphasize that most of the population is mestizo, trivializing and legally neglecting the inherent value of Indigenous ancestry (Vasconcelos, 1997; García Deister, 2014; Manrique, 2016; Wade, 2017). While Indigenous peoples have some legal recognition in these countries, the discourse of a majority mestizo nation erases Indigenous rights over ancestral lands and heritage under the illusion that all mestizos have equal rights over Pre-Hispanic cultural heritage (Endere and Ayala, 2012; Silva et al., 2022). Afrodescendant communities, who have faced historical marginalization and invisibility in Latin American countries, encounter similar challenges because national identity in these countries is so strongly tied to mestizaje (Arocha and Maya, 2008; Weltman-Cisneros and Tello, 2013; Agren, 2020). To illustrate how this misconception can permeate scholarship, a recent piece on global guidelines for aDNA research (Alpaslan-Roodenberg et al., 2021) wrongfully claimed that mestizos in many Latin American countries "embraced their Indigenous roots", hence the request to consult with Indigenous peoples in this region was "paternalistic" and "colonialist". By reproducing such harmful narratives that relegate Indigenous peoples to legacies that should be considered a matter of the past and only embedded in the present-day mestizo national identity, such statements reproduce a long history of institutional discrimination and Indigenous erasure in Latin America. While consultation with Indigenous peoples for aDNA research outside the United States is a complex subject, for which NAGPRA protocols cannot be directly applied, stating that it is not needed because Indigenous identity is well represented by the mestizo population and State institutions is a fallacy and a continuation of historical harms.

In India and other parts of South Asia, the dynamics and recognition of Indigenous identities and rights in heritage management may differ from other places discussed above and even display vast intra-regional heterogeneity but, ultimately, result in a similar undermining of Indigenous engagement and involvement that are, to our knowledge, not currently mandated in archaeological (Jamir, 2022a) or aDNA research.

In this context, from an aDNA researcher's perspective, identifying Indigenous populations that could be affiliated to individuals found in a given archaeological site or museum collection is not straightforward. In places like Mexico, where ancient empires like the Mexica invaded many territories and where multiple cultures could coexist in a single place (Mata-Míguez et al., 2012; Manzanilla, 2017), an additional layer of complexity emerges. Even if these connections can be made, challenges can surface if present-day Indigenous communities do not hold spiritual affiliation to ancestors from many generations ago (Cucina, 2013; Whittaker, 2020), if they have never been consulted about research participation before or face other more immediate challenges to their sovereignty or livelihoods (Castellanos, 2020; Hesketh, 2021; Rodriguez Aguilera, 2021).

Because of the high levels of poverty and injustice most Indigenous communities in the Global South face (Hall and Patrinos, 2012; Hall and Gandolfo, 2016), an aDNA researcher can be hesitant of whether bringing yet another issue to consult and decide upon (especially one that has never been considered before) is correct or if it is invasive or imposed.

Recommendation: Prioritize community engagement as an integral part of the research design. As outlined above and elsewhere (Wagner et al., 2020), there is no one-size-fits-all set of guidelines. Instead of foregoing the engagement process entirely because of the inherent complexity, the subject of community consultation needs to be discussed thoroughly and applied to each circumstance, while including Indigenous scholars and other stakeholders in deciding the most ethical approach for each place, regardless of whether local legislation requires consultation or Indigenous approval to carry out the project. This requires dedicated time and resources to research the history and present situation of the region and communities one wants to work with, and then preparing to engage with and involve them in horizontal discussion before, during and after the project. Integrating descendant community perspectives into research should not be seen as a burden or checkboxing an outreach step (Muller and Dortch, 2020). Instead, such discussions acknowledge and respect the richness of community-based knowledge that can additionally significantly enrich the research process (Wagner et al., 2020).

CONCLUSION

To conclude, we propose applying a "glocal" approach to aDNA research in the Global South. The glocal principle highlights the importance of assessing global-local interactions by considering the local context within a coherent global pattern. As described by (Patton, 2020), global systems must be contextually sensitive and grounded in the interaction between local and global processes. In aDNA research, this would entail applying global premises of sustainability and justice and maintaining awareness of the historical harms caused by scientific colonialism, extractivism, and other forms of exploitation of Global South nations by Global North researchers (Argüelles et al., 2022). Locally, aDNA researchers must be attuned to the implications of their research, especially regarding heritage regulation and management, knowledge and resource sharing, the development or strengthening of local expertise, involvement of Indigenous communities, and conflicts that may arise with traditional knowledge systems. More broadly, institutions in the Global North, such as funding agencies and academic promotion and tenure committees, can support efforts that prioritize community engagement by recognizing or funding this work as an integral component of the research process. Importantly, we strongly believe that for the field of aDNA to meet these ethical responsibilities, the pace must not be dictated by the growth of the field (Alpaslan-Roodenberg et al., 2021), but by prioritizing the requisite time to build and implement accountability measures. Despite the time and effort required, we find that such commitments also foster the creation and maintenance of long-term partnerships which can ultimately aid the research process and lead to a more sustainable, just and inclusive paleogenomics research field for the Global South.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

The authors would like to thank Tiatoshi Jamir, Niraj Rai, Shanti Pappu and Nizamuddin Taher for providing details on the regulation of archaeological and scientific research in India. They further acknowledge Niraj Rai, Alejandra Castillo Carbajal, Carina Uribe Díaz and Mauricio Moraga for helping with budget and shipment details for **Table 1**.

REFERENCES

Abarca Labra, V., Herrera Soto, M. J., Fuenzalida Bahamondes, N., and Valenzuela Castro, V. (2018). Cuerpos humanos de origen arqueológico. *Anales de Arqueología y Etnología* 73, 221–249.

Adame, F. (2021). Meaningful Collaborations Can End 'helicopter Research. *Nature*. doi:10.1038/d41586-021-01795-1

Agren, D. (2020). We Exist. We're Here": Afro-Mexicans Make the Census after Long Struggle for Recognition. The Guardian. Available at: https://www.theguardian.com/world/2020/mar/19/afro-mexicans-census-history-identity (Accessed February 10, 2022).

Alpaslan-Roodenberg, S., Anthony, D., Babiker, H., Bánffy, E., Booth, T., Capone, P., et al. (2021). Ethics of DNA Research on Human Remains: Five Globally Applicable Guidelines. *Nature* 599, 41–46. doi:10.1038/s41586-021-04008-x

Argüelles, J. M., Fuentes, A., and Yáñez, B. (2022). Analyzing Asymmetries and Praxis in aDNA Research: A Bioanthropological Critique. *Am. Anthropologist* 124, 130–140. doi:10.1111/aman.13692

Arocha, J., and Maya, A. (2008). "Afro-Latin American Peoples," in *A Companion to Latin American Anthropology*. Editor D. Poole John Wiley & Sons, 399–425. doi:10.1002/9781444301328.ch19

Bardill, J., Bader, A. C., Garrison, N. A., Bolnick, D. A., Raff, J. A., Walker, A., et al. (2018). Advancing the Ethics of Paleogenomics. *Science* 360, 384–385. doi:10.1126/science.aaq1131

Begay, R. L., Garrison, N. A., Sage, F., Bauer, M., Knoki-Wilson, U., Begay, D. H., et al. (2019). Weaving the Strands of Life (Iiná Bitł'ool): History of Genetic Research Involving Navajo People. *Hum. Biol.* 91, 189–208. doi:10.13110/humanbiology.91.3.04

Benn Torres, J., and Torres Colón, G. A. (2020). Genetic Ancestry: Our Stories, Our Pasts. London. Routledge Available at: https://www.routledge.com/Genetic-

Ancestry-Our-Stories-Our-Pasts/Torres-Colon/p/book/9780367566401 (Accessed February 10, 2022).

Carter, M., and Hujo, K. (2021). Universities and Social Inequalities in the Global South. Geneva:Publications | UNRISD. Available at: https://www.unrisd.org/unrisd/website/document.nsf/(httpPublications)/44021C774165371B802586F7002EA845?OpenDocument (Accessed February 10, 2022).

Castellanos, M. B. (2020). *Indigenous Dispossession: Housing and Maya Indebtedness in Mexico*. Stanford University Press. doi:10.1515/9781503614352

Champlot, S., Berthelot, C., Pruvost, M., Bennett, E. A., Grange, T., and Geigl, E.-M. (2010). An Efficient Multistrategy DNA Decontamination Procedure of PCR Reagents for Hypersensitive PCR Applications. *PLOS ONE* 5, e13042. doi:10.1371/journal.pone.0013042

Ciocca, D. R., and Delgado, G. (2017). The Reality of Scientific Research in Latin America; an Insider's Perspective. *Cell Stress Chaperones* 22, 847–852. doi:10.1007/s12192-017-0815-8

Claw, K. G., Anderson, M. Z., Begay, R. L., Tsosie, K. S., Fox, K., and Garrison, N. A. (2018). A Framework for Enhancing Ethical Genomic Research with Indigenous Communities. *Nat. Commun.* 9, 2957. doi:10.1038/s41467-018-05188-3

Claw, K. G., Dundas, N., Parrish, M. S., Begay, R. L., Teller, T. L., Garrison, N. A., et al. (2021). Perspectives on Genetic Research: Results from a Survey of Navajo Community Member. *Frontiers. Gene.* 12. Available at: https://www.frontiersin.org/article/10.3389/fgene.2021.734529 (Accessed February 10, 2022). doi:10.3389/fgene.2021.734529

Crellin, R. J., and Harris, O. J. T. (2020). Beyond Binaries. Interrogating Ancient DNA. *Arch. Dial.* 27, 37–56. doi:10.1017/S1380203820000082

Cucina, A. (2013). Ética en bioarqueología. *Temas Antropológicos. Revista Científica de Investigaciones Regionales* 35, 149–170.

Dados, N., and Connell, R. (2012). The Global South. *Contexts* 11, 12–13. doi:10.1177/1536504212436479

d'Alpoim Guedes, J., Gonzalez, S., and Rivera-Collazo, I. (2021). Resistance and Care in the Time of COVID-19: Archaeology in 2020. *Am. Anthropologist* 123, 898–915. doi:10.1111/aman.13669

DaRos, M., and Colten, R. H. (2009). A History of Caribbean Archaeology at Yale University's Peabody Museum of Natural History. *pbmb* 50, 49–62. doi:10.3374/014.050.0104

Docquier, F. (2014). *The Brain drain from Developing Countries*. IZA Institute of Labor Economics. doi:10.15185/izawol.31

Egorova, Y. (2009). De/geneticizing Caste: Population Genetic Research in South Asia. *Sci. as Cult.* 18, 417–434. doi:10.1080/09505430902806975

Endere, M. L., and Ayala, P. (2012). Normativa Legal, Recaudos Éticos Y Práctica Arqueológica: Un Estudio Comparativo De Argentina Y Chile. *Chungará (Arica)* 44, 39–57. doi:10.4067/S0717-73562012000100004

Ferreyra, M. M., Avitabile, C., Botero Álvarez, J., Haimovich Paz, F., and Urzúa, S. (2017). *At a Crossroads: Higher Education in Latin America and the Caribbean*. Washington, DC: World Bank. doi:10.1596/978-1-4648-1014-5

Fox, K., and Hawks, J. (2019). Use Ancient Remains More Wisely. *Nature* 572, 581–583. doi:10.1038/d41586-019-02516-5

Françozo, M., and Strecker, A. (2017). Caribbean Collections in European Museums and the Question of Returns. *Int. J. Cult. Property* 24, 451–477. doi:10.1017/S0940739117000248

Fulton, T. L., and Shapiro, B. (2019). Setting up an Ancient DNA Laboratory. *Methods Mol. Biol. (Clifton, N.J.),1983*, 1–13. doi:10.1007/978-1-4939-9176-1_1

Gannon, M. (2019). When Ancient DNA Gets Politicized. Smithsonian Magazine. Available at; https://www.smithsonianmag.com/history/when-ancient-dna-gets-politicized-180972639/(Accessed February 13, 2022).

García Deister, V. (2014). "Laboratory Life of the Mexican Mestizo," in *Mestizo Genomics: Race Mixture, Nation, and Science in Latin America*. Editors P. Wade, C. L. López Beltrán, E. Restrepo, and R. Ventura Santos (Durhan, North Carolina:Duke University Press). doi:10.1215/9780822376729-007

G. H. Hall and H. A. Patrinos (Editors) (2012). *Indigenous Peoples, Poverty, and Development* (Cambridge: Cambridge University Press). doi:10.1017/CBO9781139105729

Haelewaters, D., Hofmann, T. A., and Romero-Olivares, A. L. (2021). Ten Simple Rules for Global North Researchers to Stop Perpetuating Helicopter Research in the Global South. *PLOS Comput. Biol.* 17, e1009277. doi:10.1371/journal.pcbi.1009277

Hall, G., and Gandolfo, A. (2016). Poverty and Exclusion Among Indigenous Peoples: The Global Evidence. Available at: https://blogs.worldbank.org/voices/poverty-and-exclusion-among-indigenous-peoples-global-evidence (Accessed February 14, 2022).

Harmon, A. (2018). Why White Supremacists Are Chugging Milk (And Why Geneticists Are Alarmed). The New York Times. Available at: https://www.nytimes.com/2018/10/17/us/white-supremacists-science-dna.html (Accessed February 14, 2022).

Hesketh, C. (2021). Clean Development or the Development of Dispossession? the Political Economy of Wind parks in Southern Mexico. *Environ. Plann. E: Nat. Space* 2514848621991764. doi:10.1177/2514848621991764

Hudson, M., Garrison, N. A., Sterling, R., Caron, N. R., Fox, K., Yracheta, J., et al. (2020). Rights, Interests and Expectations: Indigenous Perspectives on Unrestricted Access to Genomic Data. *Nat. Rev. Genet.* 21, 377–384. doi:10.1038/s41576-020-0228-x

Instituto Nacional de Antropología e Historia. (2019). *Lineamientos generales para el manejo y resguardo de restos humanos*. Mexico: Instituto Nacional de Antropología e Historia (INAH).

Jamir, T. (2022a). "Indigenous Archaeology (Article in Press)," in *The Routledge Companion to Northeast India*. Editors J. J. P. Wouters and T. B. Subha. (Routledge).

Jamir, T. (2022b). *Personal Communication Professor Tiatoshi Jamir*.

Judd, K., and McKinnon, M. (2021). A Systematic Map of Inclusion, Equity and Diversity in Science Communication Research: Do We Practice what We Preach? *Frontiers In Communication* 6. Available at: https://www.frontiersin.org/article/10.3389/fcomm.2021.744365 (Accessed February 14, 2022).

Knapp, M., Clarke, A. C., Horsburgh, K. A., and Matisoo-Smith, E. A. (2012). Setting the Stage - Building and Working in an Ancient DNA Laboratory. *Ann. Anat.* 194, 3–6. doi:10.1016/j.aanat.2011.03.008

Lebel, J., and McLean, R. (2018). A Better Measure of Research from the Global South. *Nature* 559, 23–26. doi:10.1038/d41586-018-05581-4

Lewis-Kraus, G. (2019). Is Ancient DNA Research Revealing New Truths — or Falling into Old Traps? The New York Times. Available at: https://www.nytimes.com/2019/01/17/magazine/ancient-dna-paleogenomics.html (Accessed February 7, 2022).

Liverpool, L. (2021). Researchers from Global South Under-represented in Development Research. *Nature*. doi:10.1038/d41586-021-02549-9

Llamas, B., Valverde, G., Fehren-Schmitz, L., Weyrich, L. S., Cooper, A., and Haak, W. (2017). From the Field to the Laboratory: Controlling DNA Contamination in Human Ancient DNA Research in the High-Throughput Sequencing Era. *STAR: Sci. Technol. Archaeological Res.* 3, 1–14. doi:10.1080/20548923.2016.1258824

Llorens-Liboy, M., and Núñez, M. (2011). "Puerto Rico/Borikén/Porto Rico," in *The Routledge Handbook of Archaeological Human Remains and Legislation*. (London:Routledge). doi:10.4324/9780203838716.ch50

L. R. Manzanilla (Editor) (2017). *Multiethnicity and Migration at Teopancazco: Investigations of a Teotihuacan Neighborhood Center* (Gainesville, FL: University Press of Florida).

Maher, D., Ford, N., and Unwin, N. (2012). Priorities for Developing Countries in the Global Response to Non-communicable Diseases. *Globalization and Health* 8, 14. doi:10.1186/1744-8603-8-14

Malhi, R. S., and Bader, A. (2015). Engaging Native Americans In Genomics Research. *Am. Anthropol.* 117, 743–744. doi:10.1111/aman.12369

Manrique, L. (2016). Dreaming of a Cosmic Race: José Vasconcelos and the Politics of Race in Mexico, 1920s–1930s. *Cogent Arts & Humanities* 3, 1218316. doi:10.1080/23311983.2016.1218316

Mata-Míguez, J., Overholtzer, L., Rodríguez-Alegría, E., Kemp, B. M., and Bolnick, D. A. (2012). The Genetic Impact of aztec Imperialism: Ancient Mitochondrial DNA Evidence from Xaltocan, Mexico. *Am. J. Phys. Anthropol.* 149, 504–516. doi:10.1002/ajpa.22152

Mishra, P. (2006). Emigration and Brain Drain : Evidence from the Caribbean. *IMF*. Available at:https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Emigration-and-Brain-Drain-Evidence-From-the-Caribbean-18662 (Accessed February 13, 2022). doi:10.5089/9781451862850.001

Muller, C., and Dortch, J. (2020). "Traditional Owner Participation in Genetic Research," in *Interrogating Human Origins: Decolonisation and the Deep Human Past*. Editors M. Porr and J. Matthews. (Routledge). Available at: https://www.routledge.com/Interrogating-Human-Origins-Decolonisation-

and-the-Deep-Human-Past/Porr-Matthews/p/book/9781138300439 (Accessed February 14, 2022).

Nazario, C. R. (2014). *A Socioeconomic and Demographic Analysis with Data from the Puerto Rico Community Survey*. Puerto Rico: Inter American University of Puerto Rico, Metro Campus.

N. Marquez-Grant and L. Fibiger (Editors) (2013). *The Routledge Handbook of Archaeological Human Remains and Legislation: An International Guide to Laws and Practice in the Excavation and Treatment of Archaeological Human Remains* (London. Routledge Available at: https://www.routledge.com/The-Routledge-Handbook-of-Archaeological-Human-Remains-and-Legislation/Marquez-Grant-Fibiger/p/book/9780415859400 (Accessed February 10, 2022).

Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P. W., Ávila-Arcos, M. C., et al. (2021). Ancient DNA Analysis. *Nat. Rev. Methods Primers* 1, 1–26. doi:10.1038/s43586-020-00011-0

Pagán-Jiménez, J. (2000). *La Antropología en Puerto Rico: dicotomía de centro-periferia*. Boletín de Antropología Americana, 193.

Pagán-Jiménez, J. R., and Rodriguez Ramos, R. (2008). Sobre arqueologías de liberación en una colonia "postcolonial" (Puerto Rico). *Revista de Ciencias Sociales ISSN* 19, 8–41.

Panofsky, A., Dasgupta, K., and Iturriaga, N. (2021). How White Nationalists Mobilize Genetics: From Genetic Ancestry and Human Biodiversity to Counterscience and Metapolitics. *Am. J. Phys. Anthropol.* 175, 387–398. doi:10.1002/ajpa.24150

Pappu, S. (2022). Personal Communication. *Dr. Shanti Pappu*.

Patton, M. (2020). *Blue Marble Evaluation: Premises and Principles*. New York, NY: Guilford Press. Available at: https://www.guilford.com/books/Blue-Marble-Evaluation/Michael-Quinn-Patton/9781462541942 (Accessed February 14, 2022).

Pickrell, J. K., and Reich, D. (2014). Toward a New History and Geography of Human Genes Informed by Ancient DNA. *Trends Genet.* 30, 377–389. doi:10.1016/j.tig.2014.07.007

Rai, N. (2022). *Personal Communication*. Dr. Niraj Rai.

Reich, D. (2018). Opinion | How to Talk about 'Race' and Genetics. The New York Times. Available at: https://www.nytimes.com/2018/03/30/opinion/race-genetics.html (Accessed February 13, 2022).

Reidpath, D., and Allotey, P. (2020). Preserve Global South's Research Capacity. *Science* 368, 725. doi:10.1126/science.abc2677

Reidpath, D. D., and Allotey, P. (2019). The Problem of 'trickle-Down Science' from the Global North to the Global South. *BMJ Glob. Health* 4, e001719. doi:10.1136/bmjgh-2019-001719

Rodriguez Aguilera, M. Y. (2021). Everyday Resistances to Environmental Racism, Mestizo Geographies, and Toxicity in Oaxaca. *The Funambulist Magazine*. Available at: https://thefunambulist.net/magazine/decolonial-ecologies/everyday-resistances-to-environmental-racism-mestizo-geographies-and-toxicity-in-oaxaca (Accessed February 13, 2022).

Rodríguez López, J. (2009a). El Estado: ¿instrumento de destrucción o agente de protección del patrimonio histórico? Algunas reflexiones críticas sobre el caso de Puerto Rico. *Cuba Arqueológica* 78. Available at: https://www.academia.edu/18052069/El_Estado_instrumento_de_destrucci%C3%B3n_o_agente_de_protecci%C3%B3n_del_patrimonio_hist%C3%B3rico_Algunas_reflexiones_cr%C3%ADticas_sobre_el_caso_de_Puerto_Rico (Accessed February 14, 2022).

Rodríguez López, J. (2009b). Ley Federal NAGPRA, patrimonio arqueológico e identidad étnica en Puerto Rico. *Boletín de antropología americana*, 59–98.

Sahra. Gibbon, R. V. Santos, and Mónica. Sans (Editors) (2011). "Racial Identities, Genetic Ancestry, and Health," *South America : Argentina, Brazil, Colombia, and Uruguay/*." in *Sahra Gibbon, Ricardo Ventura Santos, and Mónica Sans. Houndmills* (Basingstoke, Hampshire: Palgrave Macmillan).

Santos Cid, A. (2022). Los recortes en el presupuesto tocan a la Escuela Nacional de Antropología e Historia. *El País México*. Available at: https://elpais.com/mexico/2022-01-03/los-recortes-en-el-presupuesto-tocan-a-la-escuela-nacional-de-antropologia-e-historia.html (Accessed February 14, 2022).

Silva, C. P., de la Fuente Castro, C., González Zarzar, T., Raghavan, M., Tonko-Huenucoy, A., Martinez, F. I., et al. (2022). The Articulation of Genomics, Mestizaje, and Indigenous Identities in Chile: A Case Study of the Social Implications of Genomic Research in Light of Current Research Practices. *Front. Genet.* doi:10.3389/fgene.2022.817318

Somel, M., Altınışık, N. E., Özer, F., and Ávila-Arcos, M. C. (2021). Collaborate Equitably in Ancient DNA Research and beyond. *Nature* 600, 37. doi:10.1038/d41586-021-03541-z

Taher, N. (2022). *Personal Communication*. Dr. Nizamuddin Taher.

TallBear, K. (2013). Genomic Articulations of Indigeneity. *Soc. Stud. Sci.* 43, 509–533. doi:10.1177/0306312713483893

Trines, S. (2018). *Economic Storm: The Crisis of Education in Puerto Rico*. WENR. Available at: https://wenr.wes.org/2018/05/economic-storm-the-crisis-of-education-in-puerto-rico (Accessed February 13, 2022).

Tsosie, K. S., Bader, A. C., Fox, K., Bolnick, D. A., Garrison, N. A., and Smith, R. W. A. (2021). Ancient-DNA Researchers Write Their Own Rules. *Nature* 600, 37. doi:10.1038/d41586-021-03542-y

Tsosie, K. S., Begay, R. L., Fox, K., and Garrison, N. A. (2020). Generations of Genomes: Advances in Paleogenomics Technology and Engagement for Indigenous People of the Americas. *Curr. Opin. Genet. Dev.* 62, 91–96. doi:10.1016/j.gde.2020.06.010

Valenzuela-Toro, A. M., and Viglino, M. (2021). How Latin American Researchers Suffer in Science. *Nature* 598, 374–375. doi:10.1038/d41586-021-02601-8

Vasconcelos, J. (1997). *The Cosmic Race/La Raza Cosmica*. 51102nd edition. Baltimore, Md: JHUP.

Viera, D., Mutize, T., and Roser Chinchilla, J. (2020). Understanding Access to Higher Education in the Last Two Decades – UNESCO-IESALC. *UNESCO Instituto Internacional para la Educación Superior en América Latina y el Caribe*. Available at: https://www.iesalc.unesco.org/en/2020/12/23/understanding-access-to-higher-education-in-the-last-two-decades/(Accessed February 13, 2022).

Wade, L. (2018). To Overcome Decades of Mistrust, a Workshop Aims to Train Indigenous Researchers to Be Their Own Genome Experts. *Sci. News*. Available at:https://www.science.org/content/article/overcome-decades-mistrust-workshop-aims-train-indigenous-researchers-be-their-own (Accessed February 20, 2022). doi:10.1126/science.aav5286

Wade, P., López-Beltrán, C., Restrepo, E., and Santos, R. V. (2015). Genomic Research, Publics and Experts in Latin America: Nation, Race and Body. *Soc. Stud. Sci.* 45, 775–796. doi:10.1177/0306312715623108

Wade, P. (2017). Mexico, Public Health, and State Genomics," in. *Degrees of Mixture, Degrees of Freedom: Genomics, Multiculturalism, and Race in Latin America*. Available at: https://read.dukeupress.edu/books/book/2313/chapter/918837/Mexico-Public-Health-and-State-Genomics (Accessed February 13, 2022).

Wagner, J. K., Colwell, C., Claw, K. G., Stone, A. C., Bolnick, D. A., Hawks, J., et al. (2020). Fostering Responsible Research on Ancient DNA. *Am. J. Hum. Genet.* 107, 183–195. doi:10.1016/j.ajhg.2020.06.017

Weinberg, B. (2011). Developing Science: Scientific Performance and Brain Drains in the Developing World. *J. Develop. Econ.* 95, 95–104. doi:10.1016/j.jdeveco.2010.05.009

Weltman-Cisneros, T., and Tello, C. D. M. (2013). Negros-Afromexicanos: Recognition and the Politics of Identity in Contemporary Mexico. undefined Available at: https://www.semanticscholar.org/paper/Negros-Afromexicanos%3A-Recognition-and-the-Politics-Weltman-Cisneros-Tello/c97c875e80bd7175b34a8ca994f48aa0a1301898 (Accessed February 14, 2022).

Whittaker, C. (2020). Aztecs Are Not Indigenous: Anthropology and the Politics of Indigeneity. *Ann. Anthropological Pract.* 44, 173–179. doi:10.1111/napa.12147

Wolinsky, H. (2019). Ancient DNA and Contemporary Politics. *EMBO Rep.* 20, e49507. doi:10.15252/embr.201949507

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

TYPE Perspective
PUBLISHED 19 August 2022
DOI 10.3389/fgene.2022.817899

OPEN ACCESS

EDITED BY
Arslan A. Zaidi,
University of Pennsylvania, United States

REVIEWED BY
Katrina Claw,
University of Colorado, Denver, United States

*CORRESPONDENCE
Katherine C. Crocker,
kcrocker@pugetsound.edu

SPECIALTY SECTION
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

RECEIVED 18 November 2021
ACCEPTED 13 July 2022
PUBLISHED 19 August 2022

CITATION
Thompson TV and Crocker KC (2022), Study the forest, not only the trees: Environmental exposures, not genomes, generate most health disparities.
Front. Genet. 13:817899.
doi: 10.3389/fgene.2022.817899

Study the forest, not only the trees: Environmental exposures, not genomes, generate most health disparities

Taylor V. Thompson[1] and Katherine C. Crocker[2]*

[1]Genetics Department, Albert Einstein College of Medicine, New York, NY, United States, [2]Biology Department, University of Puget Sound, Tacoma, WA, United States

As sequencing and analysis techniques provide increasingly detailed data at a plummeting cost, it is increasingly popular to seek the answers to medical and public health challenges in the DNA sequences of affected populations. This is methodologically attractive in its simplicity, but a genomics-only approach ignores environmentally mediated health disparities, which are well-documented at multiple national and global scales. While genetic differences exist among populations, it is unlikely that these differences overcome social and environmental factors in driving the gap in health outcomes between privileged and oppressed communities. We advocate for following the lead of communities in addressing their self-identified interests, rather than treating widespread suffering as a convenient natural experiment.

KEYWORDS

genomics, environmental exposure, health disparities, environmental racism, health equity

Introduction

Genomics is a powerful tool, and as sequencing costs plummet, increasingly popular across research and applied fields. Such popularity does not guarantee equitability in accessibility or application. Many genetic studies have fallen into the same trap as much of western medicine, focusing on the DNA of northwestern Europeans, although little biological reason for this focus has been presented. For example, between 2005 and 2020, 88% of GWAS studies presented findings using sequences from individuals of European ancestry, and 72% reported discoveries using sequences from only the United Kingdom, US, or Iceland (Mills and Rahal, 2019). Such an asymmetry in focus has resulted in calls for more equitable focus in genetic research to reduce the perception of the white European genome as the normal, healthy state. Our colleague Dr. John M. Greally has described the current status of the field of human medical genetics as facing "a balance between ignoring whole sections of the globe, and exploiting them," noting that "currently, we fall far to the side of ignoring" (JM Greally, personal communication). This disparity in degree of medical research effort between marginalized racialized people and white racialized Europeans and those descended from them mirrors national and global health disparities. However, genomics as a mechanism for addressing

population-level health disparities (that is, differences in the burden of health insults) has serious limitations.

Environment and genetics interact to produce health

Population-level health disparities are affected by both genetics and environmental exposures. A much-discussed 1975 study concerning deleterious multi-generational effects of environmental exposures focused on the descendants of individuals who survived the Dutch Hunger Winter (Stein and Susser, 1975). While the methodology and scientific conclusions of studies on this phenomenon are debated (Paneth and Susser, 1995; Schulz, 2010), the analytical approach of considering environmental determinants of health is sound. It is generally agreed that famine and other adverse environmental factors were drivers of long-term health challenges, not any genetic predisposition of Dutch individuals. Subsequent work has generally demonstrated that environment has profound effects on health (Emeny et al., 2021), yet disparities between Black and Indigenous communities and white communities are frequently assumed to be genetic in origin (effectively blaming the DNA of Black and Indigenous communities), rather than a consequence of environmental racism (Yudell et al., 2016; Borrell et al., 2021).

Global-scale evidence indicates that environmental, not genetic, explanations are strongly implicated in population-level health disparities. For example, life expectancy is a common public health measure (Roubal et al., 2021), and despite a fairly constant genetic background, it shifts dramatically based on environment (The World Bank, 2019a; The World Bank, 2019b). For example, immigrants to the United States from nations with lower life expectancies than the United States experience increased life expectancy beyond the average US-born individual, though their US-born descendants do not (Argeseanu Cunningham et al., 2008; Mehta et al., 2016; Bastian et al., 2020). This pattern is incompatible with a genetic explanation, especially in the presence of the dramatic environmental shifts that accompany immigration.

In the lands lately known as the United States, racialized minoritized people experience disproportionately high levels of direct and indirect environmental exposures (Mikati et al., 2018; Rubio et al., 2020). Acknowledging this reality can only strengthen genomic studies designed to improve health in our communities. For example, a study by Rastogi et al., in 2013 used a study design which matched cases and controls not merely concerning age, sex, and ethnicity, but also drawing from the same geographic area and hospital systems (Rastogi et al., 2013). Such a detailed study design can account for environmental exposures associated with geographic location (e.g., vehicle emissions, heavy metal exposure, water quality) and personal experiences of medical racism in a particular hospital system: such awareness is powerful for building health disparity-related genomics studies. Rigor is gained in the analysis of data beyond sequences, especially those data which are associated with environmental exposures relevant to the condition studied.

Environmental exposures are particularly relevant to health disparities because the prevalence of environmental exposures in a particular community may, if unexamined, result in the mistaken assumption that genetic predisposition is the only, or even primary, cause of observed disease. For example, the impacts of chronic community exposure to lead in drinking water may be misattributed to genetic factors if it generates similar symptoms (e.g., cardiovascular disease and reduced attention span) in multiple generations (see: Levallois et al., 2018). In fact, inaccessible resources, as well as exposures to environmental pollution and the psychological trauma associated with racism and targeted violence have detrimental short- and long-term effects on health of communities on the losing end of health disparities (Richardson and Norris, 2010; Ray, 2021). Interpersonal exposure-mediated health challenges are particularly key to address because they are "sticky" (that is, not confined to a single environment): visibly and hypervisibly racialized people experience racism in effectively all social and geographic locations (Negrete Alfaro, 2011; Linley, 2018; Niles et al., 2020).

Discussion

Environmental exposures are the results of government policies and social practices, not endogenous biological processes (Shavers and Shavers, 2006). While the function of a DNA sequence cannot be divorced from its environment, neither can the sequence itself be used as a proxy for its environment. In other words, it is impossible for any genome to produce perfect health when immersed in an acutely unhealthy environment. With the increase in computational technology and infrastructure, it is now feasible for analyses to consider not just a DNA sequence, but myriad related variables. Indeed, in failing to do so, we are systematically excising a large proportion of relevant, high-dimensional data and thereby reducing the rigor of our work (Yearby, 2020; Harrison, 2021).

Determining which environmental variables are relevant requires a degree of expertise in community experience that is frequently absent in researchers, who are frequently outsiders. However ignorant we researchers may be, community members have detailed and long-lived expertise concerning their environments and their individual and collective health: recognizing this expertise is key to addressing health disparities, especially those of racialized and marginalized communities. Unfortunately, many of these communities have

suffered extensive harm at the hands of researchers working in the name of improving health, access to resources, and human knowledge. To achieve collaboration with these communities, extra measures are necessary to safeguard community autonomy and bolster the right to withdraw consent (Zahara, 2016). Further, community-identified priorities must be recognized as the only legitimate guiding principles for research initiatives in that community. For example, the Strong Heart Study (strongheartstudy.org), an epidemiologic study of cardiovascular disease in American Indians, uses a multi-stage community consent process via which researchers apply to use tribal data for a specifically parameterized study. If consent is granted, at the conclusion of their study, the research team communicates their findings to the communities as part of their application for the communities' consent to publish that work.

In planning studies concerning the genetics of racialized and marginalized communities, researchers must exercise trust of community members, and just as importantly, skepticism of our own processes and assumptions. In other words, communities are not study systems, to have careers built on their ongoing suffering. It is very probable that our investigations will not find that more molecular research is needed. For example, if an environmental exposure is skewed towards people with presumed (see: Yudell et al., 2016) genetic backgrounds that deviate from the European "standard", exogenous (e.g., social, governmental) causes are implicated: in this case the path towards health equity lies in social and governmental change, not further genomic research. Only in the case of an evenly distributed exposure which disproportionately harms members of distinct communities, are molecular and genetics studies implicated as a potential source of solutions. And even in this case, it should be considered whether chronic environmental exposures may be more likely drivers than the genetic (Levallois et al., 2018). Genomic studies cannot overcome persistent engineered environmental effects.

In conclusion, we emphasize that mobilizing genetics research to serve oppressed and marginalized communities is not incompatible with ameliorating already-identified sources of harm. However, using genetics as a tool in moving towards a more equitable future requires substantial front-end labor to build structures which protect community autonomy and robust consent (Zahara, 2016). Many necessary changes are disincentivized in current academic and applied research contexts because they would increase the cost of research in terms of time, personnel, or resources. Nevertheless, we call for structural, contextual, and methodological changes in the way that health disparities and genomic research equity are approached. We advocate for following the lead of communities in addressing their self-identified interests, rather than treating widespread suffering as a convenient natural experiment. Until researchers recognize that health disparities are problems to be urgently addressed (often through non-research means), rather than opportunities for career advancement, health equity will remain merely aspirational.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

KCC conceived the topic of the article. KCC and TVT drafted and revised the article collaboratively; both authors approved the final version and are jointly accountable for the accuracy and integrity of this work.

Funding

KCC is funded under NIH award 3R01AG057422-05S1; TVT is funded under NIH award 5F31ES031480-02.

Acknowledgments

This paper was discussed and written on the occupied lands of the Lenape and Wappinger Nations, lately known as the New York City Metropolitan Area. Both authors are indebted to Dr. John M. Greally and Dr. Esther Choo for comments and conversation. TVT thanks Ken and Vicki Thompson, Dr. Candice Thompson for enlightening discussions and Dr. Myles Akabas for support. KC thanks Dr. Pascale R. Leroueil, Dr. Paulette Vincent-Ruz, Dr. Kishana Y. Taylor, Dr. Ana Navas-Acien, and the members of Dr. Max Liboiron's CLEAR lab for helpful comments and generative conversations. KCC also thanks Dr. Rachel E. Wilbur and Taylor S. Morriseau for important conversations about relations, harm, and obligations. Finally, KCC is particularly indebted to Da'Shaun Harrison (@dashaunLH) and Professor Dr. Arrianna Planey (@Arrianna_Planey) for their public scholarship and teaching on Twitter, and to Professors Dr. Amy Bombay and Dr. Eve Tuck for their work, respectively, on intergenerational trauma and communication of harm.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Argeseanu Cunningham, S., Ruben, J. D., and Venkat Narayan, K. M. (2008). Health of foreign-born people in the United States: A review. *Health Place* 14, 623–635. doi:10.1016/j.healthplace.2007.12.002

Bastian, B., Tejada Vera, B., and Arias, E. (2020). *Mortality trends in the United States, 1900-2018*. (United States: National Center for Health Statistics).

Borrell, L. N., Elhawary, J. R., Fuentes-Afflick, E., Witonsky, J., Bhakta, N., Wu, A. H. B., et al. (2021). Race and genetic ancestry in medicine—a time for reckoning with racism. *Obstetrical Gynecol. Surv.* 76, 395–397. doi:10.1097/01.ogx.0000767204.20020.0e

Emeny, R. T., Carpenter, D. O., and Lawrence, D. A. (2021). Health disparities: Intracellular consequences of social determinants of health. *Toxicol. Appl. Pharmacol.* 416, 115444. doi:10.1016/j.taap.2021.115444

Harrison, D. S. L. (2021). *Belly of the beast*. (Berkeley, California, USA: North Atlantic Books).

Levallois, P., Barn, P., Valcke, M., Gauvin, D., and Kosatsky, T. (2018). Public health consequences of lead in drinking water. *Curr. Environ. Health Rep.* 5, 255–262. doi:10.1007/s40572-018-0193-0

Linley, J. L. (2018). Racism here, racism there, racism everywhere: The racial realities of minoritized peer socialization agents at a historically white institution. *J. Coll. Student Dev.* 59, 21–36. doi:10.1353/csd.2018.0002

Mehta, N. K., Elo, I. T., Engelman, M., Lauderdale, D. S., and Kestenbaum, B. M. (2016). Life expectancy among U.S.-born and foreign-born older adults in the United States: Estimates from linked social security and medicare data. *Demography* 53, 1109–1134. doi:10.1007/s13524-016-0488-4

Mikati, I., Benson, A. F., Luben, T. J., Sacks, J. D., and Richmond-Bryant, J. (2018). Disparities in distribution of particulate matter emission sources by race and poverty status. *Am. J. Public Health* 108, 480–485. doi:10.2105/AJPH.2017.304297

Mills, M. C., and Rahal, C. (2019). A scientometric review of genome-wide association studies. *Commun. Biol.* 2, 9. doi:10.1038/s42003-018-0261-x

Negrete Alfaro, M. A. (2011). *When the invisible punishing machine is everywhere... The mechanism of social control (mass incarceration, institutionalized racism, slavery and repression) in the USA shapes the individual as well as the social space*. (Cambridge: Massachusetts Institute of Technology).

Niles, P. M., Asiodu, I. V., Crear-Perry, J., Julian, Z., Lyndon, A., Mclemore, M. R., et al. (2020). Reflecting on equity in perinatal care during a pandemic. *Health Equity* 4, 330–333. doi:10.1089/heq.2020.0022

Paneth, N., and Susser, M. (1995). Early origin of coronary heart disease (the "Barker hypothesis"). *BMJ* 310, 411–412. doi:10.1136/bmj.310.6977.411

Rastogi, D., Suzuki, M., and Greally, J. M. (2013). Differential epigenome-wide DNA methylation patterns in childhood obesity-associated asthma. *Sci. Rep.* 3, 2164. doi:10.1038/srep02164

Ray, K. (2021). In the name of racial justice: Why bioethics should care about environmental toxins. *Hastings Cent. Rep.* 51, 23–26. doi:10.1002/hast.1251

Richardson, L. D., and Norris, M. (2010). Access to health and health care: How race and ethnicity matter. *Mt. Sinai J. Med.* 77, 166–177. doi:10.1002/msj.20174

Roubal, A. M., Pollock, E. A., Gennuso, K. P., Blomme, C. K., and Givens, M. L. (2021). Comparative methodologic and practical considerations for life expectancy as a public health mortality measure. *Public Health Rep.* 137, 255–262. doi:10.1177/0033354921999407

Rubio, R., Grineski, S., Collins, T., and Morales, D. X. (2020). Ancestry-based intracategorical injustices in carcinogenic air pollution exposures in the United States. *Soc. Nat. Resour.* 33, 987–1005. doi:10.1080/08941920.2019.1708521

Schulz, L. C. (2010). The Dutch Hunger Winter and the developmental origins of health and disease. *Proc. Natl. Acad. Sci. U. S. A.* 107, 16757–16758. doi:10.1073/pnas.1012911107

Shavers, V. L., and Shavers, B. S. (2006). Racism and health inequity among Americans. *J. Natl. Med. Assoc.* 98, 386–396.

Stein, Z., and Susser, M. (1975). The Dutch famine, 1944–1945, and the reproductive process. I. Effects on six indices at birth. *Pediatr. Res.* 9, 70–76. doi:10.1203/00006450-197502000-00003

The World Bank. (2019a). "Life expectancy at birth, total (years) - north America". Available at: https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=XU, (Accessed 1 Mar, 2022).

The World Bank. (2019b). "Life expectancy at birth, total (years) - sub-Saharan Africa". Available at: https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=ZG, (Accessed 1 Mar, 2022).

Yearby, R. (2020). Structural racism and health disparities: Reconfiguring the social determinants of health framework to include the root cause. *J. Law Med. Ethics* 48, 518–526. doi:10.1177/1073110520958876

Yudell, M., Roberts, D., Desalle, R., and Tishkoff, S. (2016). SCIENCE AND SOCIETY. Taking race out of human genetics. *Science* 351, 564–565. doi:10.1126/science.aac4951

Zahara, A. (2016). Ethnographic refusal: A how to guide discard studies Available at: https://discardstudies.com/2016/08/08/ethnographic-refusal-a-how-to-guide/, (Accessed 8 August, 2016).

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 22 August 2022
DOI 10.3389/fgene.2022.949422

OPEN ACCESS

EDITED BY
Arslan A. Zaidi,
University of Pennsylvania, United States

REVIEWED BY
Shweta Ramdas,
Azim Premji University, India
Samantha Baxter,
Broad Institute, United States

*CORRESPONDENCE
Jennifer L. Young,
youngjl@stanford.edu

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

RECEIVED 20 May 2022
ACCEPTED 11 July 2022
PUBLISHED 22 August 2022

CITATION
Young JL, Halley MC, Anguiano B, Fernandez L, Bernstein JA, Wheeler MT and Tabor HK (2022), Beyond race: Recruitment of diverse participants in clinical genomics research for rare disease.
Front. Genet. 13:949422.
doi: 10.3389/fgene.2022.949422

Beyond race: Recruitment of diverse participants in clinical genomics research for rare disease

Jennifer L. Young[1,2*†], Meghan C. Halley[1†], Beatriz Anguiano[3], Liliana Fernandez[4], Jonathan A. Bernstein[5], Matthew T. Wheeler[6] and Holly K. Tabor[1,7]
Undiagnosed Diseases Network Consortium

[1]Stanford Center for Biomedical Ethics, Stanford University School of Medicine, Stanford, CA, United States, [2]Center for Genetic Medicine, Northwestern Feinberg School of Medicine, Chicago, IL, United States, [3]Human Genetics and Genetic Counseling, Stanford University School of Medicine, Stanford, CA, United States, [4]Stanford Center for Undiagnosed Diseases, Stanford University, Stanford, CA, United States, [5]Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, United States, [6]Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, United States, [7]Department of Medicine, Stanford University, Stanford, CA, United States

Purpose: Despite recent attention to increasing diversity in clinical genomics research, researchers still struggle to recruit participants from varied sociodemographic backgrounds. We examined the experiences of parents from diverse backgrounds with enrolling their children in clinical genomics research on rare diseases. We explored the barriers and facilitators parents encountered and possible impacts of sociodemographic factors on their access to research.

Methods: We utilized semi-structured interviews with parents of children participating in the Undiagnosed Diseases Network. Interview data were analyzed using comparative content analysis.

Results: We interviewed 13 Hispanic, 11 non-Hispanic White, four Asian, and two biracial parents. Participants discussed different pathways to clinical genomics research for rare disease as well as how sociodemographic factors shaped families' access. Themes focused on variation in: 1) reliance on providers to access research; 2) cultural norms around health communication; 3) the role of social capital in streamlining access; and 4) the importance of language-concordant research engagement.

Conclusion: Our findings suggest that variables beyond race/ethnicity may influence access in clinical genomics research. Future efforts to diversify research participation should consider utilizing varied recruitment strategies to reach participants with diverse sociodemographic characteristics.

KEYWORDS

rare disease, equity, genome sequencing, exome sequencing, pediatrics

Introduction

Access to clinical genomics research for people of diverse sociodemographic identities is essential for achieving equity in the distribution of benefits from the knowledge gained. Unequal access to opportunities to participate in research is not in adherence with the principle of justice, which would require that all members of society benefit equitably from scientific advancement (Green et al., 2020) Furthermore, to the extent that diverse genetic ancestries correlate with sociodemographic categories of race and ethnicity, the lack of diverse participants in genomics research reduces the broad applicability of findings and limits classification of rare variants in individuals from underrepresented groups (Bonkowsky et al., 2018; Landry and Rehm, 2018)

Calls have been made to increase recruitment of racially/ethnically diverse participants in clinical genomics research since the vast majority of participants have been of European descent (Roberts et al., 2018; Ceyhan-Birsoy et al., 2019; Fatumo et al., 2022) The issue of equity in access has been particularly pointed in the context of rare diseases, for which patients face additional challenges related to care coordination and access to knowledgeable specialists (Splinter et al., 2018; Walley et al., 2018) Over the last decade, research has advanced the use of genomic sequencing for gene discovery and diagnosis of rare diseases, with the potential to improve access to genetic diagnosis for rare disease patients (Spillmann et al., 2017; Posey et al., 2019) Indeed, as many patients face insurance and other barriers to accessing genomic sequencing in clinical care, research has become a source of access to diagnostic tools such as sequencing for patients (Delikurt et al., 2015) As a result, in the context of rare disease, failure to reduce inequities in access to research also may contribute to health disparities in access to genetic diagnosis for rare disease patients.

Prior research provides some insights into barriers to research participation that disproportionately impact certain communities. A number of sociodemographic factors are known to shape access to health research, including cultural norms and beliefs related to health and illness, lack of education, financial resources and health insurance (Lee et al., 2019; McGuire et al., 2020; Fatumo et al., 2022) Especially for historically excluded or exploited groups, a lack of trust in research–and in healthcare institutions more broadly–has been reported as a central reason for declining to participate in research. There is evidence that this is particularly problematic in clinical genomics research, where certain racial/ethnic groups report concerns about privacy and whether genomic information will be used against them by the government, healthcare system, or law enforcement–issues that disproportionately affect certain groups (Amendola et al., 2018; Claw et al., 2018; Passmore et al., 2019)

Though many barriers to clinical genomics research have been identified, the experiences of diverse participants who have successfully enrolled in clinical genomics research remain less well understood. These individuals' perspectives are valuable insomuch as they may speak to not only potential barriers, but also the facilitators to research access. Further, the literature on barriers to access in clinical genomics research has relied heavily on the lens of race/ethnicity, with less emphasis placed on the ways in which other sociodemographic factors may shape access.

To address these gaps, we conducted interviews with parents of diverse racial/ethnic backgrounds to examine their experiences of enrolling their children in clinical genomics research, including barriers and facilitators they encountered, and how various sociodemographic factors shaped their access to research. We focused on access to clinical genomics research for diverse patients with rare genetic diseases, including barriers and facilitators to identification of research opportunities, recruitment for and enrollment in studies, and participant retention.

Materials and methods

Study setting

We conducted this study in collaboration with the Stanford University clinical site of the Undiagnosed Diseases Network (UDN) (Reuter et al., 2018) The UDN is a research consortium developed to advance the science of rare disease discovery and diagnosis through a case-based approach (Gahl et al., 2016) Any individual may apply to the UDN, though a referring provider letter is required. Enrollment is based on multiple criteria, including the presence of an undiagnosed condition despite thorough evaluation by a health care provider, the presence of at least one objective finding, and willingness to consent to, travel for (when necessary), and participate in the recommended clinical, research, and genetic evaluation (Ramoni et al., 2017)

As of December 2021, the UDN had evaluated over 1,500 participants and diagnosed 505 of those individuals (Network, 2021) Applicants to the UDN are (>80%) White and Non-Hispanic. The network reports no difference in rate of acceptance for different racial/ethnic groups among those who complete the application (Walley et al., 2018) At the Stanford University site 49.5% of enrolled pediatric participants identified as Non-Hispanic White, 26.3% as White Hispanic, 11.8% as Asian, 8.2% as Multiracial, and 1.5% as Black or African American.

The UDN study is approved by a central institutional review board at the National Institutes of Health and is registered at clinicaltrials.gov (NCT02450851) (Splinter et al., 2018) This study was also approved by the Stanford University Institutional Review Board.

Participant recruitment

A UDN clinical site coordinator provided the lead researcher on the study team with contact information for parents of UDN participants who previously agreed to be contacted for future research. We utilized quota sampling (Bernard, 2006) to ensure racial/ethnic diversity. We focused recruitment on Asian American and Hispanic families, the two largest racial/ethnic minority groups at the study site. We intentionally recruited non-Hispanic White participants for one-third of our sample as a comparison group. A Spanish-English bilingual clinical site coordinator and bilingual researchers helped to recruit Spanish-speaking parents. Researchers contacted potential participants through phone and email. Individuals were eligible to participate if they were the parent or legal guardian of a current UDN participant and if their primary language was either English or Spanish.

Data collection

Parents who consented to participate in the study completed a single in-depth, semi-structured interview in either English or Spanish, lasting between one and 2 hours. After reviewing the literature, the study team developed the interview guide and pilot tested it with parents with children with undiagnosed or rare diseases. The final interview guide included questions regarding the participant's sociodemographic characteristics, family structure, their child's diagnostic odyssey, and experiences before, during and after participating in the UDN, including barriers to and facilitators of access to research. Interviews were audio-recorded and transcribed verbatim, translated from Spanish to English (when necessary), and de-identified for analysis.

Data analysis

The research team analyzed interview data using Dedoose software (Dedoose (9.0.17), 2021) Three experienced qualitative researchers (JLY, MCH, HKT) utilized inductive and deductive approaches to conduct a comparative content analysis of the data (Miles et al., 2018) First the research team reviewed the transcripts to define deductive codes related to broad content area (e.g., "family," "healthcare experiences," "genetic testing experiences"). We conducted repeated interrater reliability testing in the application of these codes until the average pooled Cohen's kappa reached $\kappa > 0.8$, indicating excellent agreement (Miles et al., 2018) We then applied the deductive broad codes to all transcripts. Drawing on inductive techniques from Grounded Theory (Strauss et al., 1998), we then used these codes to iteratively explore potential mediating factors driving similarities and differences in participants' experiences accessing research by various sociodemographic characteristics, including race/ethnicity, primary language, education, and income. This process included attention to both expressed (emic) differences in access as well as observed (etic) differences within and across different subsets of our sample (Strauss et al., 1998)

TABLE 1 Sample characteristics.

	N	%
Parent gender		
Female	27	93.3
Male	3	10.0
Parent Race/Ethnicity		
Hispanic (any race)	13	43.3
White (Non-Hispanic)	11	36.6
Asian-American (Non-Hispanic)	4	13.3
More than one race/ethnicity	2	6.6
Parent Preferred Language		
English	20	66.6
Spanish	10	33.3
Household Combined Income		
< $50,000	7	23.3
$50,000 - $100,000	7	23.3
$100,001 - $150,000	9	30.0
$150,001 - $200,000	4	13.3
> $200,000	3	10.0
Highest Education Completed		
No school	1	3.3
Elementary school	1	3.3
Some high school	5	16.7
High school	4	13.3
Some college	3	10.0
College	7	23.3
Graduate degree	9	30.0
Number of Children in UDN		
One	23	76.7
Two	6	20.0
Three	1	3.3
Child(ren)'s Diagnostic Status		
Diagnosed	13	43.3
Undiagnosed	14	46.7
Emerging/Candidate Diagnosis	3	10.0
Total All Parent Participants	**30**	**100**

Bold values are the our participants were all parents, this is how we describe them in the table in comparison to any data about their children.

Results

Participant characteristics

We completed interviews with one parent from 30 unique families. Twenty participants completed interviews in English and 10 participants completed interviews in Spanish. Parents self-identified as Asian ($n = 4$), non-Hispanic White ($n = 11$),

Hispanic (n = 14), and multiracial (n = 2). Participating parents were predominantly female (n = 27, 93.3%), and diverse in terms of income and education. See Table 1 for full sample characteristics.

Barriers and facilitators to clinical genomic research in rare disease

The results of our analysis are organized into four themes regarding aspects of participants' experiences with clinical genomics research that suggest potential barriers and/or facilitators to access. While the first two themes relate to participant race/ethnicity, the second two highlight the extent to which additional sociodemographic characteristics, beyond race/ethnicity, also may shape families' access to clinical genomics research. Specifically, themes highlight variation in: 1) reliance on providers to access research; 2) cultural norms around health communication; 3) the role of social capital in streamlining access; and 4) the importance of language-concordant research engagement. Below we describe these four themes, the connections across themes, and provide supporting quotes.

Theme 1: Reliance on providers to access research

The first theme focuses on the role of healthcare providers in facilitating access to research. Hispanic parent participants more commonly reported relying on and trusting providers to help their child and to facilitate access to clinical genomic research. This trust existed across income and education levels, as well as English language proficiency. These parents shared a distinct description of gratitude towards providers for their persistence in searching for a diagnosis, finding therapies, and helping families navigate the medical system. This was especially relevant for parents who reported struggling to understand complex medical information and did not feel confident in their ability to provide their child with appropriate care. For example:

> *There's a lot of need and more necessity in our culture because there's a [lack of knowledge]. We don't understand that there are different diseases that we don't know about. So it is up to the professionals not to give up. If God gave them that knowledge to research, to study, it is so they can help more people live a life that perhaps is not normal, but is better. (P29, Mother, Hispanic)*

In responding to a question about what advice they would offer to other families trying to access research, participants also described the importance of a close relationship with providers.

> *What I would say is . . . believe in the doctors. Believing that there are people who are interested in helping others, in this case the doctors who are interested in our children (P20, Mother, Hispanic)*

One parent described how this relationship could be especially important for families with language barriers.

> *Especially for the families that don't speak English, I would say really to have a close relationship with your primary provider–primary care. And that you feel that you could tell them anything, I think that is key. I feel a lot of parents they know something's wrong but they don't know how to take care of it. When my daughter was sick I didn't know, I was young, I was naïve, I didn't know. (P04, Mother, Hispanic)*

The degree of trust in providers prominent in the narrative of Hispanic participants was distinct from that of White and Asian parents. Non-Hispanic families more commonly described compromised trust in providers, and/or the healthcare system as a whole.

> *We didn't trust doctors for a very long time. (P02, Mother, Asian)*

> *That was definitely an eye opener because I trusted our provider so much . . . so much. And I think it broke a lot of trust for both my husband and myself. (P15, Mother, White)*

As discussed further below, White and Asian parents in our study population were more likely to report higher incomes and education levels, and this access to social capital appears to have facilitated more direct access to clinical genomics research despite distrust.

Theme 2: Cultural norms around health communication

The second theme focuses on variation in reported cultural norms around health communication. More frequently in Asian participants, communication–or lack of communication–about a child's illness was described as shaping the ways in which they accessed research opportunities. Specifically, three out of the four Asian participants independently described a tendency for individuals in the Asian community to conceal or avoid discussing issues related to illness or disability, even among close family members. For example, P01 shared that, "it would be definitely harder for [an Asian family] to speak and talk about the situation," and described how she herself has struggled to communicate with her parents about her

child's condition. She also shared a story of a friend and fellow Asian mother who also had only told a few people about her child's rare disease.

Though participants narratives focused on health communication within families, they also identified implications of these cultural norms for access to research and suggested strategies to overcome this barrier. For example, P02 pointed out that putting information online is important to families who are less likely to discuss their child's condition with others.

> *Get an [Asian] family to talk about a medical issue, it is not going to happen. They will hide, hide, hide. They will not be as open with sharing the data, but they are research oriented. Best thing is to put it online, they will Google the condition. Word-of-mouth will not work with 60–70% of Asians. (P02, Mother, Asian)*

This challenge was not reported among other participants, who described more open communication among immediate and extended family as well as other support networks. In contrast to P01, a White mother (P16), said that "we don't keep [our child's] care or diagnosis or, you know, journey or anything a secret from anyone."

Theme 3: Role of social capital in streamlining access

The third theme focused on the role of social capital–including existing social networks and the ability to find and navigate such connections–in facilitating some parents' ability to directly access clinical genomics research. This theme was prominent in English-speaking participants of varying racial/ethnic backgrounds with high levels of education and family income, and notably absent among participants with lower levels of education and income.

For example, one parent who worked in the sciences independently identified the opportunity to participate in clinical genomics research by talking to physicians at her work about her child.

> *When I talked to physicians at my work, and there's one in particular . . . I told her who I had reached out to and she's like that's perfect, go to the UDN. (P07, mother, White)*

Another parent identified the opportunity to participate in research through a non-profit organization to which her family had given charitable donations in the past.

> *It was my husband who actually connected with [non-profit organization] and then started donating to that network–I think that's how we got connected with the UDN and then got enrolled. (P13, mother, Asian)*

Another parent found out about the opportunity to participate in clinical genomics research while he was attending a fundraising event, during which he made the acquaintance of someone directly involved in clinical genomics research.

> *I basically met [a doctor's wife] at an event. At the conclusion of the event I just met her serendipitously and we were talking about kids, and I mentioned S (my daughter), and we talked a little and she said, well, you got to meet [my husband] because [he] found out last week about an $80 million National Institute of Health funded study for people with undiagnosed medical conditions. (P19, father, White)*

Other parents focused on their own efforts and abilities in independently gathering information and navigating potential research opportunities. These individuals all reported having a master's degree or higher, demonstrated a high level of genetic literacy, and described doing extensive research.

> *And then [the providers] went to [genetic testing] panels . . . but now we're going to do whole exome, and then finally, like we were not offered whole genome but that's when I was like, well, we've had two rounds of this, like I want to go somewhere else. And so that's when we went to the UDN. (P14, Father, White)*

> *I was kind of doing my own research and I was like, wait, I think I've heard of this before. And I went back and (our geneticist) told me about this! And so then I called from there. I really focused on it for the first few years, like just pursuing and pursuing, and since that had been my focus we had a lot of data already which helped I think get us in. (P06, Mother, Biracial)*

Not all parents in the highest income brackets and highest level of education groups discussed leveraging social capital to directly access research. However, no families without such resources described doing so and some even noted how difficult this might be.

> *I see [other] parents and they look for resources, it is like, I just don't have the bandwidth. I'm just so exhausted. (P05, Mother, Hispanic)*

Theme 4: Importance of language-concordant research engagement

The fourth theme focuses on the importance of language-concordant communication in facilitating recruitment from diverse patients who had limited English proficiency (LEP). These parents stressed the value of Spanish language communication in accessing research. One LEP parent described

how having a Spanish-speaking provider serve as a navigator during the application process for the UDN helped access this resource:

> *The geneticist, she's a very good person, told me: "Look, here's the form, fill it out." At that time there was a speech therapist with my daughter and she helped me to fill the form out and entered her over the computer. (P27, Mother, Hispanic)*

Another parent highlighted the importance of Spanish language services particularly in the context of genetics, due to the complexity of the conversations required. She said that although she does not always require interpretation services:

> *. . .when I go to genetics, I ask if they can give me an interpreter because the doctor talks using numbers and codes that I don't understand. But normally I do the appointments alone. (P24, Spanish-speaking mother)*

Though this participant is referring to the clinical context, her comment also has implications for clinical genomics research.

Retention of participants is an equally important aspect of access to research and Spanish-speaking parents spoke very positively about the communication and support they received from a Spanish-speaking research coordinator working for the Stanford UDN.

> *We're happy with (Spanish-speaking UDN staff) who has been talking with us, she's telling us every time they find something or if they look at something or if they have not found anything yet, but they keep letting us know. They make sure we don't think they forget her. (P25, Spanish-speaking mother)*

This experience among LEP parents stood in contrast to those described by many of the English-speaking parents. Indeed, for English-speaking parents, a lack of consistent and clear communication was their primary complaint regarding their experiences with clinical genomics research.

> *I would love for [the UDN] to communicate more often . . . I don't want to feel like a second afterthought. (P08, Mother, White)*

While we cannot conclude that Spanish language was the only driver of increased satisfaction, given the differences by language across our sample, this suggests at least a reasonable hypothesis. Providing language-concordant support in navigating study participation appeared to improve the quality of communication between participants and researchers beyond even what was experienced by English-speaking participants.

Discussion

Our results describe potential barriers and facilitators to accessing and participating in clinical genomics research for parents of children with rare diseases from sociodemographically diverse backgrounds, including relationships with providers, cultural norms around communication, and access to social capital and language-concordant research resources. Parents' narratives highlighted sociodemographic factors–including income, education, cultural norms, and language proficiency–that may play a role in shaping access, either separately or in addition to race and ethnicity.

In this study, Hispanic participants more commonly stressed the importance of having a close and trusting relationship with a provider in facilitating access to research opportunities. In contrast, Asian participants more frequently described cultural barriers to communication but also identified online resources as a potential avenue identifying research opportunities. However, it was only the subset of participants who had high income and education that reported being able to leverage social capital to facilitate access to research. The role of cultural norms, language, education, and income illustrated in our themes highlight the importance of examining multiple sociodemographic characteristics in addition to race/ethnicity when considering barriers and facilitators to clinical genomics research.

Our results align with literature on the role of communication in clinical research, and the extent to which the quality and consistency of communication may influence patient identification and enrollment practices (Sae-Hau et al., 2021) Language-concordant engagement has been shown to facilitate navigation of clinical genetic services for individuals with limited English proficiency (Pacyna et al., 2021; de Leon et al., 2022) Our study suggests that resources for communication may be equally important in genomics research, where concepts and terminology are very complex and non-native English speakers who do not typically need an interpreter may struggle with communication. Given that nearly one in 10 adults in the United States have limited English proficiency, language-concordant research services are an essential tool for inclusive research (Ryan, 2013)

Our findings also intersect in unexpected ways with the literature on the role of trust in recruitment of historically marginalized communities in clinical genomics research (George et al., 2013; Claw et al., 2018; Kraft et al., 2018; Lee et al., 2019; Armstrong and Ritchie, 2022) While this literature has emphasized trust–and specifically mistrust–as a barrier to clinical genomics research participation, Hispanic parents in our study expressed *more* trust in providers to identify and refer them to clinical genomics research than other parents. Prior research suggests that those who are less familiar with the bureaucracy of clinical genomics research may rely more heavily on providers and researchers to facilitate access (Levine et al., 2001) Research also points to the important role of providers in helping patients access clinical genetics services and could be expanded to understand how these providers may also facilitate participation in clinical genomics research (Chou et al., 2021) More research is needed to understand the barriers that these

providers themselves may face in equitably referring patients for clinical genomic research.

Implications for clinical genomics research

Recently, leaders in the scientific community have pledged to address structural racism in biomedical research through efforts focused on diversifying the genetics workforce and research participant populations (Kaiser, 2021), changes to publication policies (Brothers et al., 2021), and to research funding (Health NIo, 2021) In addition to these efforts, our findings point to the need for an intersectional approach to recruitment and retention. Rather than using race/ethnicity as a proxy for other sociodemographic characteristics, clinical genomics researchers should plan to systematically collect a broader range of variables associated with structural inequities, such as income, education, and language proficiency. Tracking this data throughout study design, implementation, and analysis is critical to promoting greater inclusion of participants from underrepresented populations that face compounding barriers to research (Bentley et al., 2017)

Recruitment strategies must address heterogeneity in access to research. Participants in our sample varied in their pathways to accessing clinical genomics research. Some relied heavily on their clinical providers to identify opportunities and facilitate access, while others leveraged social connections and/or their own research to identify these opportunities independently. Investigators developing recruitment strategies could leverage these varied approaches to reach diverse patient populations. Online resources such as clinicaltrials.gov may be a key resource for some families but less accessible to those with limited education or English-language proficiency. For other families, tools such as patient navigation may be more helpful (Fouad et al., 2016; Uveges et al., 2018) In clinical trials research, patient navigators have been utilized to meet individual needs and address barriers or concerns of participants enrolled (Ghebre et al., 2014; Uveges et al., 2018) These programs also have the potential to serve as a conduit between clinical providers and researchers. However, research is needed to effectively translate these models into clinical genomics research for specific populations.

Clinical genomic researchers must exercise caution to avoid privileging access to research participation to only those who have the resources and skills to independently identify research opportunities. Funding agencies, such as the National Institutes of Health, could promote or incentivize the adoption of research recruitment and retention efforts that are compatible with the language and cultural diversity present in the populations.

Limitations

While the racial/ethnic diversity of our sample is a strength, our study did not include the important perspectives of other groups such as Black patients and families due to small numbers at the study site as well as lack of response to recruitment efforts. The Asian American sample was also small and included people who identified as East Asian and South Asian, thus our conclusions about this subgroup may not be broadly generalizable to the many heterogeneous cultures represented by the term "Asian." In addition, our participants include those who were able to successfully access clinical genomics research. Although we explicitly probed for insights into reaching diverse communities more broadly, directly soliciting the experiences of those who were not able to access research, or who chose not to participate, remains critical. This is a small sample from a single clinical genomics study and further research with larger samples is needed to examine how diverse patient characteristics relate to research access, recruitment, and retention outside of the context of rare disease and clinical genomics research for diagnostic purposes.

Conclusion

This study suggests multiple sociodeomographic factors–including and in addition to race/ethnicity–may related to barriers and facilitators to clinical genomics research. Future research must look beyond race/ethnicity variables to better incorporate factors such as education, income, social networks, and cultural norms in planning recruitment and retention efforts to expand accessibility. Research on the ethical, legal, and social implications of genetic and genomic research often recruits convenience samples from larger genomics studies and must avoid the pitfalls and biases of convenience sampling. Facilitating research participation for all people, not just those who have sociodemographic advantages, will ensure equitable access to the direct benefits of research such as receiving a potential genetic diagnosis, as well as the indirect downstream benefits of generalizable knowledge in genomic research to communities and patients that have historically been excluded.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Stanford University Institutional Review Board. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

JY: Conceptualization, Methodology, Investigation, Formal Analysis, Writing—original draft preparation, review and editing MH: Conceptualization, Methodology, Writing, Formal Analysis, Writing—Review and editing BA: Formal Analysis, Writing—Review and editing LF: Resources, Writing—Review and editing JB: Resources, Writing—Review and editing MW: Resources, Writing—Review and editing HT: Conceptualization, Supervision, Formal Analysis, Writing—Review and editing.

Funding

Research reported in this manuscript was supported in part by the NIH Common Fund, through the Office of Strategic Coordination/Office of the NIH Director under Award Number U01HG010218, and by a grant from the Stanford Maternal & Child Health Research Institute, award number 1233211-102-JHAJL. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Stanford Maternal & Child Health Research Institute. Effort for JY was supported by T32HG008953 and 19SFRN34830054. Effort for MCH was supported by K01HG011341 and UL1TR003142. Effort for LF, JAB, and MTW was supported by U01HG010218. Effort for HKT was provided by UL1TR003142.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Amendola, L. M., Robinson, J. O., Hart, R., Biswas, S., Lee, K., Bernhardt, B. A., et al. (2018). Why patients decline genomic sequencing studies: Experiences from the CSER consortium. *J. Genet. Couns.* 27 (5), 1220–1227. doi:10.1007/s10897-018-0243-7

Armstrong, K., and Ritchie, C. (2022). Research participation in marginalized communities — overcoming barriers. *N. Engl. J. Med.* 386 (3), 203–205. doi:10.1056/NEJMp2115621

Bentley, A. R., Callier, S., and Rotimi, C. N. (2017). Diversity and inclusion in genomic research: Why the uneven progress? *J. Community Genet.* 8 (4), 255–266. doi:10.1007/s12687-017-0316-6

Bernard, H. R. (2006). *Research methods in anthropology : Qualitative and quantitative approaches*. Lanham: AltaMira Press.

Bonkowsky, J. L., Wilkes, J., Bardsley, T., Urbik, V. M., and Stoddard, G. (2018). Association of diagnosis of leukodystrophy with race and ethnicity among pediatric and adolescent patients. *JAMA Netw. Open* 1 (7), e185031. doi:10.1001/jamanetworkopen.2018.5031

Brothers, K. B., Bennett, R. L., and Cho, M. K. (2021). Taking an antiracist posture in scientific publications in human genetics and genomics. *Genet. Med.* 23 (6), 1004–1007. doi:10.1038/s41436-021-01109-w

Ceyhan-Birsoy, O., Murry, J. B., Machini, K., Lebo, M. S., Yu, T. W., Fayer, S., et al. (2019). Interpretation of genomic sequencing results in healthy and ill newborns: Results from the BabySeq project. *Am. J. Hum. Genet.* 104 (1), 76–93. doi:10.1016/j.ajhg.2018.11.016

Chou, A. F., Duncan, A. R., Hallford, G., Kelley, D. M., and Dean, L. W. (2021). Barriers and strategies to integrate medical genetics and primary care in underserved populations: A scoping review. *J. Community Genet.* 12 (3), 291–309. doi:10.1007/s12687-021-00508-5

Claw, K. G., Anderson, M. Z., Begay, R. L., Tsosie, K. S., Fox, K., Garrison, N. A., et al. (2018). A framework for enhancing ethical genomic research with Indigenous communities. *Nat. Commun.* 9 (1), 2957. doi:10.1038/s41467-018-05188-3

de Leon, A., McCarthy Veach, P., Bro, D., and LeRoy, B. S. (2022). Spanish language concordance in genetic counseling sessions in the United States: Counselor experiences and perceptions of its effects on processes and outcomes. *J. Genet. Couns.* 31, 188–205. doi:10.1002/jgc4.1472

Dedoose (9.0.17) (2021). *A web application for managing, analyzing, and presenting qualitative and mixed method research data*. Los Angeles, CA: SocioCultural Research Consultants, LLC. [computer program].

Delikurt, T., Williamson, G. R., Anastasiadou, V., and Skirton, H. (2015). A systematic review of factors that act as barriers to patient referral to genetic services. *Eur. J. Hum. Genet.* 23 (6), 739–745. doi:10.1038/ejhg.2014.180

Fatumo, S., Chikowore, T., Choudhury, A., Ayub, M., Martin, A. R., Kuchenbaecker, K., et al. (2022). A roadmap to increase diversity in genomic studies. *Nat. Med.* 28 (2), 243–250. doi:10.1038/s41591-021-01672-4

Fouad, M. N., Acemgil, A., Bae, S., Forero, A., Lisovicz, N., Martin, M. Y., et al. (2016). Patient navigation as a model to increase participation of african Americans in cancer clinical trials. *J. Oncol. Pract.* 12 (6), 556–563. doi:10.1200/JOP.2015.008946

Gahl, W. A., Mulvihill, J. J., Toro, C., Markello, T. C., Wise, A. L., Ramoni, R. B., et al. (2016). The NIH undiagnosed diseases program and network: Applications to modern medicine. *Mol. Genet. Metab.* 117 (4), 393–400. doi:10.1016/j.ymgme.2016.01.007

George, S., Duran, N., and Norris, K. (2013). A systematic review of barriers and facilitators to minority research participation among african Americans, latinos, asian Americans, and pacific islanders. *Am. J. Public Health* 104 (2), e16–e31. doi:10.2105/AJPH.2013.301706

Ghebre, R. G., Jones, L. A., Wenzel, J. A., Martin, M. Y., Durant, R. W., Ford, J. G., et al. (2014). State-of-the-science of patient navigation as a strategy for enhancing minority clinical trial accrual. *Cancer* 120 (S7), 1122–1130. doi:10.1002/cncr.28570

Green, E. D., Gunter, C., Biesecker, L. G., Di Francesco, V., Easter, C. L., Feingold, E. A., et al. (2020). Strategic vision for improving human health at the Forefront of Genomics. *Nature* 586 (7831), 683–692. doi:10.1038/s41586-020-2817-4

Health NIo (2021). *Ending structural racism*. Available at: https://www.nih.gov/ending-structural-racism (Accessed.

Kaiser, J. (2021). NIH apologizes for 'structural racism, ' pledges change. *Science* 371 (6533), 977. doi:10.1126/science.371.6533.977

Kraft, S. A., Cho, M. K., Gillespie, K., Halley, M., Varsava, N., Ormond, K. E., et al. (2018). Beyond consent: Building trusting relationships with diverse populations in precision medicine research. *Am. J. Bioeth.* 18 (4), 3–20. doi:10.1080/15265161.2018.1431322

Landry, L. G., and Rehm, H. L. (2018). Association of racial/ethnic categories with the ability of genetic tests to detect a cause of cardiomyopathy. *JAMA Cardiol.* 3 (4), 341–345. doi:10.1001/jamacardio.2017.5333

Lee, S. S., Fullerton, S. M., Saperstein, A., and Shim, J. K. (2019). Ethics of inclusion: Cultivate trust in precision medicine. *Sci. (New York, NY)* 364 (6444), 941–942. doi:10.1126/science.aaw8299

Levine, R. A., Levine, S. E., and Schnell, B. (2001). "Improve the women": Mass schooling, female literacy, and worldwide social change. *Harv. Educ. Rev.* 71 (1), 1–50. doi:10.17763/haer.71.1.154550622x3225u7

McGuire, A. L., Gabriel, S., Tishkoff, S. A., Wonkam, A., Chakravarti, A., Furlong, E. E. M., et al. (2020). The road ahead in genetics and genomics. *Nat. Rev. Genet.* 21 (10), 581–596. doi:10.1038/s41576-020-0272-6

Miles, M. B. (2018). *Qualitative data analysis : A methods sourcebook*. Editors A. M. Huberman and J. Saldaña. Fourth edition.

Network, U. D. (2021). *UDN facts and figures*. Accessed https://undiagnosed.hms.harvard.edu/about-us/facts-and-figures/.

Pacyna, J. E., Shaibi, G. Q., Lee, A., Byrne, J. O., Cuellar, I., Sutton, E. J., et al. (2021). Increasing access to individualized medicine: A matched-cohort study examining latino participant experiences of genomic screening. *Genet. Med.* 23 (5), 934–941. doi:10.1038/s41436-020-01079-5

Passmore, S. R., Jamison, A. M., Hancock, G. R., Abdelwadoud, M., Mullins, C. D., Rogers, T. B., et al. (2019). I'm a little more trusting": Components of trustworthiness in the decision to participate in genomics research for african Americans. *Public health genomics* 22 (5-6), 215–226. doi:10.1159/000505271

Posey, J. E., O'Donnell-Luria, A. H., Chong, J. X., Harel, T., Jhangiani, S. N., Coban Akdemir, Z. H., et al. (2019). Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet. Med.* 21 (4), 798–812. doi:10.1038/s41436-018-0408-7

Ramoni, R. B., Mulvihill, J. J., Adams, D. R., Allard, P., Ashley, E. A., Bernstein, J. A., et al. (2017). The undiagnosed diseases network: Accelerating discovery about health and disease. *Am. J. Hum. Genet.* 100 (2), 185–192. doi:10.1016/j.ajhg.2017.01.006

Reuter, C. M., Brimble, E., DeFilippo, C., Dries, A. M., Enns, G. M., Ashley, E. A., et al. (2018). A new approach to rare diseases of children: The undiagnosed diseases network. *J. Pediatr.* 196, 291–297. e292. doi:10.1016/j.jpeds.2017.12.029

Roberts, J. S., Robinson, J. O., Diamond, P. M., Bharadwaj, A., Christensen, K. D., Lee, K. B., et al. (2018). Patient understanding of, satisfaction with, and perceived utility of whole-genome sequencing: Findings from the MedSeq project. *Genet. Med.* 20 (9), 1069–1076. doi:10.1038/gim.2017.223

Ryan, C. (2013). *Language use in the United States: 2011*. U.S. Census Bureau.

Sae-Hau, M., Disare, K., Michaels, M., Gentile, A., Szumita, L., Treiman, K., et al. (2021). Overcoming barriers to clinical trial participation: Outcomes of a national clinical trial matching and navigation service for patients with a blood cancer. *JCO Oncol. Pract.* 17 (12), e1866–e1878. doi:10.1200/OP.20.01068

Spillmann, R. C., McConkie-Rosell, A., Pena, L., Jiang, Y. H., Schoch, K., Walley, N. M., et al. (2017). A window into living with an undiagnosed disease: Illness narratives from the undiagnosed diseases network. *Orphanet J. Rare Dis.* 12 (1), 71. doi:10.1186/s13023-017-0623-3

Splinter, K., Adams, D. R., Bacino, C. A., Bellen, H. J., Bernstein, J. A., Cheatle-Jarvela, A. M., et al. (2018). Effect of genetic diagnosis on patients with previously undiagnosed disease. *N. Engl. J. Med.* 379 (22), 2131–2139. doi:10.1056/NEJMoa1714458

Strauss, A., Corbin, J. M., and Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: SAGE Publications

Uveges, M. K., Lansey, D. G., Mbah, O., Gray, T., Sherden, L., Wenzel, J., et al. (2018). Patient navigation and clinical trial participation: A randomized controlled trial design. *Contemp. Clin. Trials Commun.* 12, 98–102. doi:10.1016/j.conctc.2018.09.003

Walley, N. M., Pena, L. D. M., Hooper, S. R., Cope, H., Jiang, Y. H., McConkie-Rosell, A., et al. (2018). Characteristics of undiagnosed diseases network applicants: Implications for referring providers. *BMC Health Serv. Res.* 18 (1), 652. doi:10.1186/s12913-018-3458-2

frontiers | Frontiers in Genetics

TYPE Policy and Practice Reviews
PUBLISHED 13 October 2022
DOI 10.3389/fgene.2022.901515

OPEN ACCESS

EDITED BY
Tina Lasisi,
University of Southern California,
United States

REVIEWED BY
Graciela Cabana,
The University of Tennessee,
United States
Xiaoye Jin,
Guizhou Medical University, China

*CORRESPONDENCE
Maria Corazon A. De Ungria,
madeungria@up.edu.ph
Jae Joseph Russell B. Rodriguez,
jae_rodriguez@eva.mpg.de

SPECIALTY SECTION
This article was submitted to ELSI in Science and Genetics, a section of the journal Frontiers in Genetics

RECEIVED 22 March 2022
ACCEPTED 21 September 2022
PUBLISHED 13 October 2022

CITATION
Rodriguez JJRB, Cuales JMD, Herrera MJB, Zubiri LAM, Muallil RN, Ishmael AI, Jimenez EB, Stoneking M and De Ungria MCA (2022), Ethical challenges in genetic research among Philippine Indigenous Peoples: Insights from fieldwork in Zamboanga and the Sulu Archipelago.
Front. Genet. 13:901515.
doi: 10.3389/fgene.2022.901515

Ethical challenges in genetic research among Philippine Indigenous Peoples: Insights from fieldwork in Zamboanga and the Sulu Archipelago

Jae Joseph Russell B. Rodriguez[1,2,3]*, John Meldwin D. Cuales[1], Michael James B. Herrera[4], Louward Allen M. Zubiri[5], Richard N. Muallil[6], Altan I. Ishmael[7], Edlyn B. Jimenez[8], Mark Stoneking[3,9] and Maria Corazon A. De Ungria[1,10]*

[1]DNA Analysis Laboratory, Natural Sciences Research Institute, College of Science, University of the Philippines Diliman, Quezon City, Philippines, [2]Genetic and Molecular Biology Division, Institute of Biological Sciences, College of Arts and Sciences, University of the Philippines Los Baños, Laguna, Philippines, [3]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, [4]Archaeological Studies Program, University of the Philippines Diliman, Quezon City, Philippines, [5]Independent Researcher, Malabon City, Philippines, [6]Office of Continuing Education and Extension Services, Mindanao State University—Tawi-Tawi College of Technology and Oceanography, Tawi-Tawi, Philippines, [7]Sama Studies Center, Mindanao State University—Tawi-Tawi College of Technology and Oceanography, Tawi-Tawi, Philippines, [8]National Institutes of Health, University of the Philippines Manila, Manila City, Philippines, [9]Universite Lyon 1, CNRS, Laboratoire de Biometrie et Biologie Evolutive, Villeurbanne, France, [10]Program on Biodiversity, Ethnicity, and Forensics, Philippine Genome Center, University of the Philippines, Quezon City, Philippines

The Philippines, with the recent discovery of an archaic hominin in Luzon and an extensive ethnolinguistic diversity of more than 100 Indigenous peoples, is crucial to understanding human evolution and population history in Island Southeast Asia. Advances in DNA sequencing technologies enable the rapid generation of genomic data to robustly address questions about origins, relatedness, and population movements. With the increased genetic sampling in the country, especially by international scientists, it is vital to revisit ethical rules and guidelines relevant to conducting research among Indigenous peoples. Our team led fieldwork expeditions between 2019 and February 2020 in Zamboanga and the Sulu Archipelago, a chain of islands connecting the Mindanao and Borneo landmasses. The trips concluded with a collection of 2,149 DNA samples from 104 field sites. We present our fieldwork experience among the mostly sea-oriented Sama-Bajaw and Tausug-speaking communities and propose recommendations to address the ethical challenges

Abbreviations: AO, Administrative order; BARMM, Bangsamoro Autonomous Region in Muslim Mindanao; CIOMS, Council for International Organizations of Medical Sciences; IKSPs, Indigenous Knowledge Systems and Practices; MIPA, Ministry of Indigenous Peoples Affairs; MoA, Memorandum of Agreement; MoU, Memorandum of Understanding; MSU-TCTO, Mindanao State University—Tawi-Tawi College of Technology and Oceanography; MTA, Material Transfer Agreement; NCIP, National Commission on Indigenous Peoples; NEGHR, National Ethical Guidelines for Health and Health-Related Research; OSCC, Office for Southern Cultural Communities; PHREB, Philippine Health Research Ethics Board; RECs, Research Ethics Committees.

of conducting such research. This work contributes toward building an enabling research environment in the Philippines that respects the rights and autonomy of Indigenous peoples, who are the rightful owners of their DNA and all genetic information contained therein.

KEYWORDS

population genetics, Philippine Indigenous peoples, research ethics, Zamboanga, Sulu Archipelago, Sama, Tausug

1 Introduction

The Philippines, an archipelagic nation in Island Southeast Asia, has figured in at least two important migration events in the region, namely the human settlement of Sunda and Sahul about 40,000 years ago (YA) (O' Connell and Allen, 2004) and the spread of Austronesian speaking farmers from Taiwan around 4,000–5,000 YA (Blust, 1995; Bellwood, 1997; Gray et al., 2009). Today the country is inhabited by more than 100 ethnolinguistic groups exhibiting cultural and phenotypic diversity (United Nations, 2007). The so-called "Negrito" phenotype possessed by certain groups (Barrows, 1910) and a "sea-nomadic" lifestyle adopted by some coastal communities (Nimmo, 2001) are of particular anthropological interest. These unique populations, along with the discovery of a new hominin species, *Homo luzonensis,* in the largest Philippine island of Luzon (Détroit et al., 2019), increasingly generate attention from scholars interested in how human evolution and prehistory unfolded in this part of the globe.

Of notable historical, linguistic, and anthropological importance is the Sulu Archipelago, a chain of islands stretching in a northeast-southwest direction between the Zamboanga peninsula of Mindanao Island in the Philippines and northeast Borneo (Figure 1). This area is divided into the three provinces of Basilan, Sulu, and Tawi-Tawi, which are part of the Bangsamoro Autonomous Region in Muslim Mindanao (BARMM; see list of abbreviations and acronyms used in this article). Isabela City in Basilan Island and Zamboanga City at the southwestern extremity of Mindanao Island belong to Region IX, one of the country's 16 administrative regions. Jolo Island, near the center, was the seat of the Sultanate of Sulu, a dominant maritime power that emerged 600 years ago following the introduction of Islam

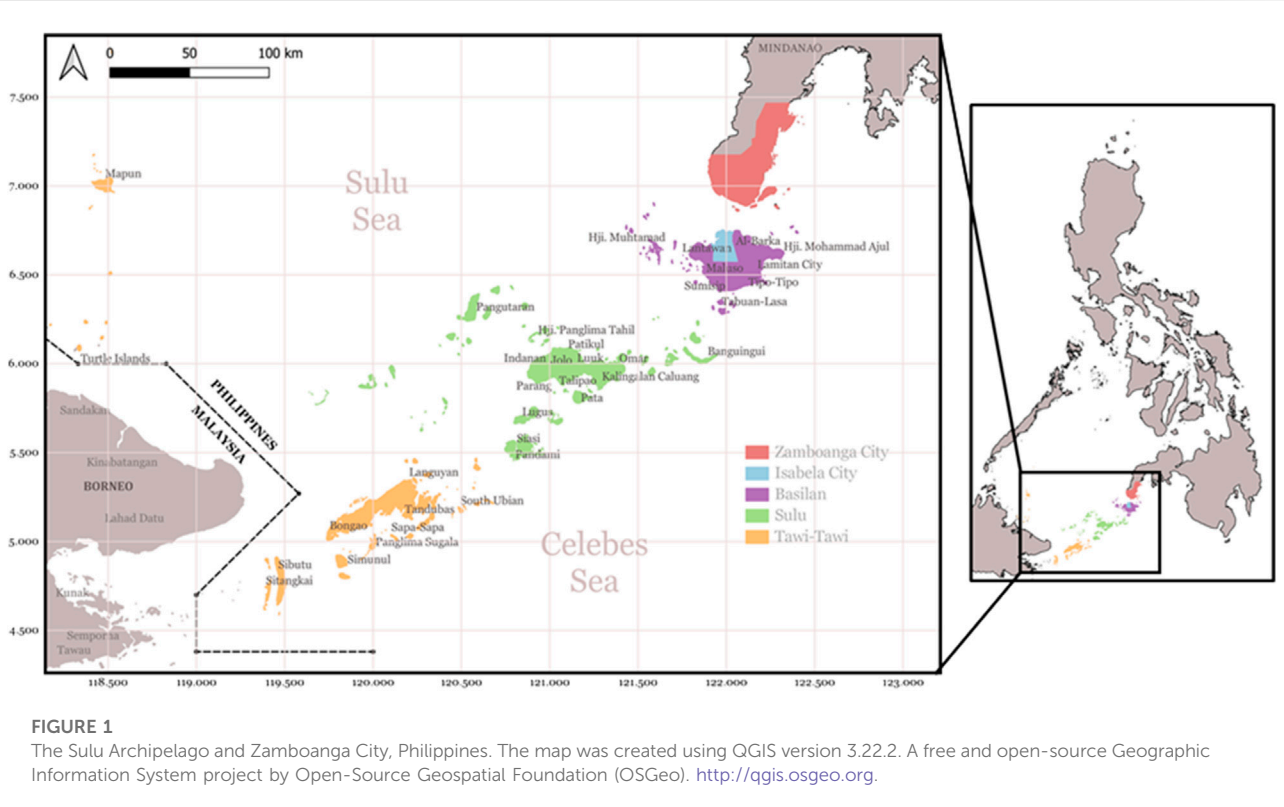


FIGURE 1
The Sulu Archipelago and Zamboanga City, Philippines. The map was created using QGIS version 3.22.2. A free and open-source Geographic Information System project by Open-Source Geospatial Foundation (OSGeo). http://qgis.osgeo.org.

(Warren, 1985). Jolo is considered the homeland of the Tausug people, whose diaspora has spread to most parts of the archipelago. Seven Indigenous languages are spoken in the island chain as defined in the Ethnologue (Eberhard et al., 2021): Yakan, Sama Bangingi, Sama Pangutaran, Central Sama, Southern Sama, Jama Mapun, and Tausug[1]. Except for Tausug, these languages are grouped as Sama-Bajaw and are classified under Greater Barito languages (Pallesen, 1985; Blust, 2007). Although the Tausug language is currently the region's lingua franca, the islands are populated by economically diverse groups of Sama speakers whose dialects are associated with certain villages, islands, or municipalities (Sather, 1997; Nimmo, 2001). The Yakan and Jama Mapun reside in Basilan and Mapun Islands, respectively, both groups having unmistakable cultural and linguistic connections to the Sama. The Sama Dilaut, more commonly known as the Badjao, live in coastal stilt houses throughout the region. As former boat-dwellers, they are often described as "sea nomads" or "sea gypsies" (Sather, 1997; Nimmo, 2001; Bellina et al., 2021).

DNA is a powerful tool for investigating questions about origins, prehistoric movements, and biological affinities. Earlier genetic studies of Philippine groups focused on uniparental lineages, shedding light on maternal and paternal histories (Tabbada et al., 2010; Delfin et al., 2011; Gunnarsdóttir et al., 2011; Delfin et al., 2014). With the development of single nucleotide polymorphism (SNP) microarrays and advanced sequencing technologies, genome-wide and high-coverage whole-genome sequence data promise more robust investigations of past demographic events and genetic adaptations (Reich et al., 2011; Lipson et al., 2014; Pagani et al., 2016; Skoglund et al., 2016; Jinam et al., 2017, GenomeAsia100K Consortium, 2019). Recent genetic studies spanning the entire country put forward far-reaching generalizations about the number, origins, and timeframe of dispersals into the Philippines (Larena et al., 2021a) and levels of archaic introgression from Denisovans (Larena et al., 2021b). Questions were raised concerning the ethical compliance of the researchers (National Commission on Indigenous Peoples, 2015; National Commission on Indigenous Peoples, 2021a; Philippine Genome Center, 2021; Rochmyaningsih, 2022). The long-term social ramifications of such alleged violations of ethical compliance are yet to be discerned. Nonetheless, past genetic studies lacking a sound ethical framework have resulted in several social harms, such as violation of individual rights, stigmatization, and general distrust of genomic research by Indigenous peoples (TallBear, 2007; Sterling, 2011; Chennells and Steenkamp, 2018). Given the growing interest from international researchers keen to study Island Southeast Asia, there is an urgent need to revisit Philippine policies, formulate clear guidelines where gaps remain, and require the compliance of all stakeholders to recognize and protect the rights of Indigenous peoples.

From 2019 to February 2020, our team conducted sampling campaigns in local communities in Zamboanga City and the Sulu Archipelago to investigate population genetic history and adaptations to a marine-oriented lifestyle. In this paper, we describe our best field practices and provide an overview of the conduct of genetic research among Philippine Indigenous peoples in compliance with existing national regulations. We intend this to be a helpful resource in reviewing research and ethical policies governing human genetics research in the Philippines.

2 Ethical guidelines for human genetic research in the Philippines

Human research in the Philippines requires compliance with national ethical rules and guidelines (Figure 2). The Republic Act No. 10532, or the Philippine National Health Research System Act of 2013, mandates that all health and health-related research should adhere to the guidelines of the Philippine Health Research Ethics Board (PHREB) (Philippine National Health Research System, 2013). PHREB adopts the WHO definition of health which is "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity" (World Health Organization, 2006, p.1). Thus, studies involving human subjects conducted in the country, including human population genetic research, are within its purview, even if not necessarily medical in scope. They must undergo "must undergo ethical review and clearance before implementation to ensure the safety, dignity, and well-being of research participants" (Department of Science and Technology, 2012, p.1). Such reviews can only be conducted by research ethics committees (RECs) accredited by PHREB.

The involvement of Indigenous peoples in research adds another level of review and monitoring by institutions mandated to protect their rights and interests, primarily the National Commission on Indigenous Peoples (NCIP), established through Republic Act 8371 or the Indigenous Peoples Rights Act of 1997. The NCIP defines Indigenous peoples: as "indigenous on account of their descent from the populations which inhabited the country, at the time of conquest or colonization, or at the time of inroads of non-indigenous religions and cultures, or the establishment of present state boundaries, who retain some or all of their own social, economic, cultural and political institutions" (National Commission on Indigenous Peoples, 1997, p.3). The NCIP has published a list of Indigenous peoples it recognizes (National

1 We apply the names of ethnic groups to the names of the languages, consistent with the Ethnologue (Eberhard et al., 2021). Whereas these may deviate from the actual terms used by locals, they are, in our opinion, more neutral to the varying degrees of usage and preference by the locals. For example, there are at least two other names for the language spoken by the Tausug, namely *Bahasa Sūg* and *Sinug*. When "Tausug language" or "speaking Tausug" appears in the manuscript, it is a shorthand for "the language spoken by the Tausug people".

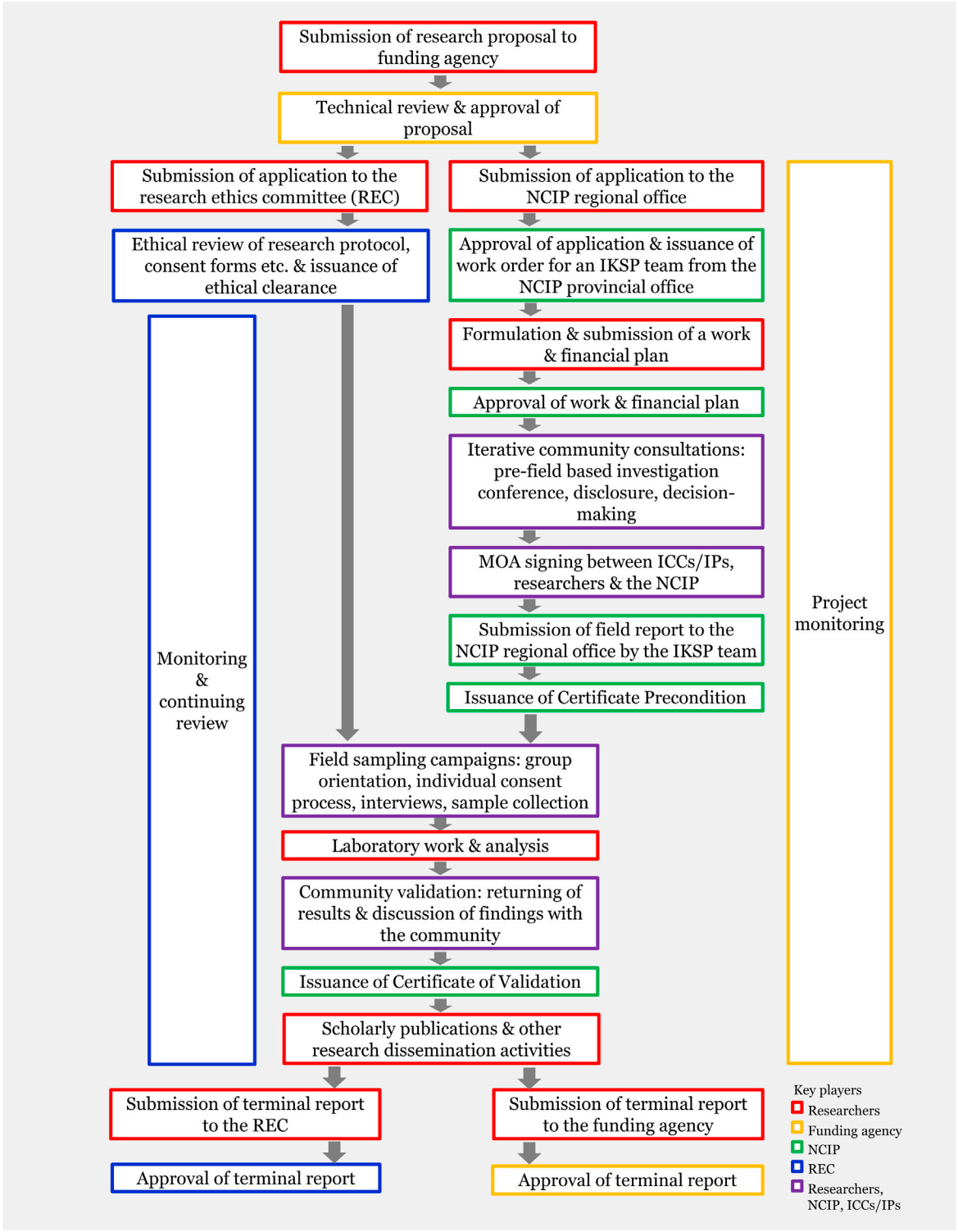


FIGURE 2
Stages of basic academic research among Philippine Indigenous cultural communities/Indigenous peoples (ICCs/IPs).

Commission on Indigenous Peoples, 2021b). Within BARMM, the Ministry of Indigenous Peoples Affairs (MIPA) performs similar functions and superseded the Office for Southern Cultural Communities (OSCC) after the promulgation of the Bangsamoro Organic Law in 2019.

The NCIP released two administrative orders in 2012. NCIP Administrative Order No. 1, known as "The Indigenous Knowledge Systems and Practices (IKSPs) and Customary Laws (CLs) Research and Documentation Guidelines of 2012" provides guidelines for academic research and community-initiated studies that could be used for policy formulations and/or implementation of NCIP mandates (National Commission on Indigenous Peoples, 2012a). NCIP Administrative Order No. 3 or "The Revised Guidelines on Free and Prior Informed Consent (FPIC) and related processes" applies to field-based investigations needed to ensure the protection of the rights "to ancestral domains, social justice, and human rights, self-governance and empowerment, and cultural integrity" (p.1) in projects aimed at commercializing Indigenous products and knowledge and those that could affect ancestral domains of the communities (National Commission on Indigenous Peoples, 2012b). In human genetic research, there is a need to distinguish basic academic research from studies with potential commercial gain (e.g., drug discovery), given the uncertainties in understanding bioprospecting, data and sample ownership, and the wealth of new information contained in individual genomes. Both Administrative Orders require forming an IKSP team, approving field plans, conducting iterative community consultations, and signing a Memorandum of Agreement (MoA) between the Indigenous groups, the researchers, and the NCIP. The MoA requires the researchers to return to the field sites, discuss the research findings, and consult the communities before publishing the study results. This is known as the research validation phase, which ensures the participants are among the first to learn about the results of the study.

The requirement for compliance with the NCIP AOs by researchers involved in studies with Indigenous groups was upheld in a Memorandum of Understanding (MoU) between NCIP and PHREB (Philippine Health Research Ethics Board, 2016), which aimed to reinforce each agency's mandate. Under this MoU, researchers must obtain clearances from the NCIP and a PHREB-accredited Level 2 REC for all studies involving Indigenous peoples. This cooperation resulted in the requirement for research to include 1) an iterative and documented process of community consultations; 2) the use of informed consent forms that are understandable to all participants and preferably translated into the language of the Indigenous peoples; 3) biobanking and data sharing policies that include provisions for removal of samples/data; and 4) data privacy requirements following the Data Privacy Act of 2012.

PHREB also published the National Ethical Guidelines for Health and Health-Related Research (NEGHHR) in 2017 with a specific section on Indigenous peoples (Philippine Health Research Ethics Board, 2017). This section reiterates the requirement to obtain a clearance from the NCIP and discusses pertinent issues such as cultural sensitivity, vulnerability, and benefit-sharing and ownership. The NEGHHR also has a provision for the transfer of custody of biological samples to foreign institutions, which should follow a Material Transfer Agreement (MTA) that complies with all applicable international and Philippine regulations. The MTA must define the responsibilities of foreign researchers, including identifying a local counterpart researcher following the CIOMS guideline on collaborative partnership and capacity-building for research (Council for International Organizations of Medical Sciences, 2016).

3 Methodology: Regulatory compliance, social preparations, and fieldwork

This human genetic research with linguistics and animal archaeogenetic components comprises an interdisciplinary project investigating the history and adaptations of the peoples of Zamboanga and the Sulu Archipelago. The proposal underwent technical review by the Philippine Commission on Higher Education and the Max Planck Institute for Evolutionary Anthropology as the Ph.D. thesis of JJRBR. Ethical approval to conduct the study was granted by the University of the Philippines Manila Research Ethics Board (UPMREB 2018-453-01) and the Ethics Council of the Max Planck Society (Application No: 2021-22). UPMREB continuously monitors the project implementation and reviews amendments to the approved protocol.

Following NCIP AO No. 1, the fieldworkers and the IKSP team organized iterative disclosure processes, decision-making, and MoA signing with Sama/Sama Dilaut and Yakan communities in Basilan and Zamboanga City. Upon reviewing the field report, the NCIP Region IX office issued the Certificate Precondition, which officially signifies the full compliance of the researchers to NCIP requirements. We also applied for and received similar clearances from the OSCC in Tawi-Tawi and Sulu. While the Tausug is not among the Indigenous groups recognized by the NCIP or the OSCC, we nonetheless conducted consultations at the village level and committed to a similar validation process, as we recognized the value of following similar procedures for all communities in the study. On all field trips, the team conducted community activities in the presence of NCIP or OSCC representatives.

A collaborative partnership was formalized with the Mindanao State University–Tawi-Tawi College of Technology

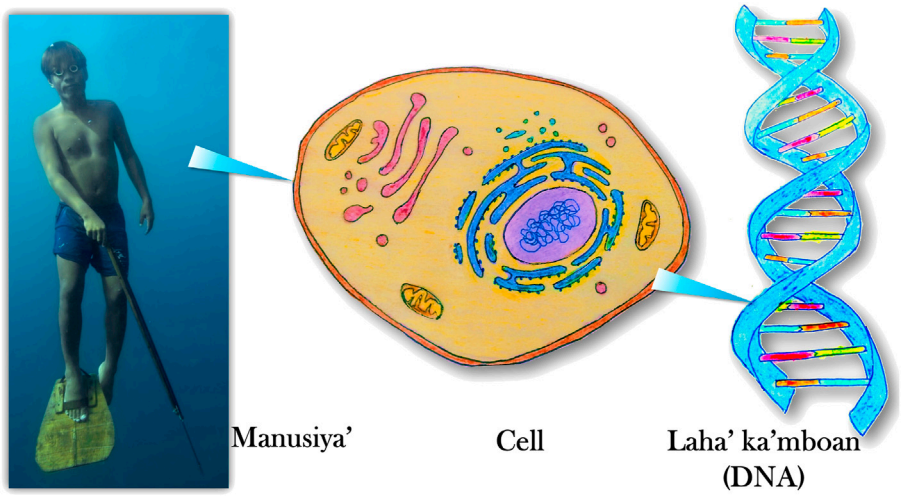


FIGURE 3
One of the visual aids used during group orientations explains how humans (*manusiya'*) are made up of cells containing DNA (*laha' ka'mboan* lit. Blood of ancestors). Illustrated by Jacob Barbosa. Photograph from Jacob Maentz.

and Oceanography (MSU-TCTO), which serves as the local counterpart university in Tawi-Tawi. The MSU-TCTO researchers assisted the research team in the field and translated consent forms and agreements into Sama and Tausug.

Because the region is recognized as a site of violent conflict (International Alert, 2020), security precautions were observed by requesting the police to accompany the research team to some sites, particularly in the Sulu province. The police officers were primarily local Tausug or Sama and were often members of the communities. In our impression, they did not influence the consent process, and their presence indicated peace and order in the vicinity. With the local government divided into provinces, cities, municipalities, and *barangays* (small administrative districts corresponding to villages) the team approached governors, mayors, chairpersons, or their representatives during each courtesy call, where the research was explained, and assistance in reaching the communities was sought. On-site, the team consulted with the village leaders before meeting with the locals.

Sample collections proceeded in communities that signified consent. Fieldworkers discussed the study objectives and procedures with prospective participants in the local language using visual aids (Figure 3) with assistance from locals who were native speakers of Sama or Tausug. Individual consent was signified by signing or placing thumbprints on approved forms for participation in this study, for potential secondary use of samples, and a 15-year provision for sample storage. Donors were 18 years or older, except for some children who were part of family trios. Minors completed the assent forms that accompanied the parental consent forms. The field team also collected participants' age, birthplace, group affinity, birthplace, and group affinity for their parents and grandparents (if known), diet, lifestyle, oral health practices, height, and weight. During the interview, the research team reconstructed the pedigrees of a participant's immediate and extended relatives. Finally, a 2 ml saliva sample was provided by each volunteer.

The research team conducted five fieldwork expeditions from 2019 to February 2020, which involved multiple trips across islands on sea vessels. It culminated in a collection of 2,149 DNA samples from 104 villages, the most extensive set of samples from any Philippine region. At the time of this writing, sample processing and data analysis are ongoing simultaneously at the Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, and the DNA Analysis Laboratory and the Philippine Genome Center at the University of the Philippines Diliman.

4 Challenges and recommendations

4.1 Difficulties in communicating the value of genetic research to the communities

Communication at a basic level was a challenge from the outset. Despite being part of the Philippines, the Sulu Archipelago has a degree of distinction from the rest of the country, given its history of staunch resistance to colonial rule, a Muslim majority, and a long-standing clamor for autonomy or separation from the Philippine state. In 2019, the Bangsamoro Organic Law was ratified, formally installing a regional parliamentary government within the unitary and presidential national government (Official Gazette, 2018). The populations of

the Sulu Archipelago predominantly speak languages other than the national language (Filipino), namely Sama and Tausug. Particularly in Sama Dilaut communities, many individuals had not received formal schooling and thus were not able to read and converse in Filipino.

To facilitate fieldwork, JJRBR and JMDC took on the challenge of learning to converse in the Sama language. Assisted by local partners, they explained the study, responded to questions, and conducted the interviews during community consultations in 104 field sites. During conversations in the field, it was evident that the concept of DNA was foreign to the worldview of many locals. However, they could grasp the concepts of ancestry and inheritance. Blood is viewed as a carrier of biological traits in the Filipino psyche, i.e., the gene (Tan, 2008). The idiom "it is in the blood" indicates that certain characteristics run in the family. One co-author, A.I.I. coined the term *laha' ka'mboan* meaning "blood of the ancestors" as an equivalent term for DNA. In doing so, the DNA was likened to a bond that connects someone alive to one's ancestors, and by extension, to others to whom they may be related "by blood". That such an entity is present in all body parts made the participants appreciate why their saliva samples had to be collected. This demonstrates how scientists can communicate their research better by translating technical questions into more relatable and profound inquiries humans have always asked: where we come from and how we are related to one another.

4.2 Potential misconceptions and inconsistencies between genetic results with indigenous peoples' oral histories

Communicating research findings accurately but in culturally acceptable ways will be challenging for the next phase of the study. Prior to publication, genetic results must undergo an output validation. The research team must return to inform and co-interpret the study findings with the communities (Figure 2). Gaps in understanding the limits of genomic inference occasionally raised inquiries on whether DNA can be used to test one's group membership, like the "blood quantum" concept (McKay, 2021), or determine one's descent from a prominent historical figure. Moreover, potential inconsistencies between genetic findings and folk narratives may prompt disagreeing sentiments. Previous research in other disciplines is instructive in this regard. For example, historical linguistics postulated the timing of the arrival of the Tausug in Jolo (Pallesen, 1985), which countered some local perceptions of autochthony in the region.

To avoid confusion when explaining the genetic findings, the team included researchers who teach in MSU-TCTO and belong to the Sama and Tausug communities. They are involved in discussions during data generation and analysis. In addition, the preparation for the validation phase will involve more local partners from MSU-TCTO, NCIP, the former OSCC, and MIPA. Field materials will be consulted with community leaders before they are used to resolve possible disputes with oral histories while maintaining the scientific integrity of the findings.

4.3 Defining different populations and the need for researchers to understand the cultural/historical basis for these groupings

In common practice, fieldwork in population genetic research involves recruiting participants of self-reported unadmixed ancestry who can at least ascribe that their four grandparents belong to the same ethnic group or reside in the general location. The availability of distinct groupings is desirable as most genetic analyses require that data be grouped into discrete samples representing defined ancestry groups, geographic regions, or languages of interest. However, in this study, grouping individuals into discrete clusters is challenging due to two factors: 1) the continuous nature of biological diversity and 2) the complexities of ethnonyms in the Sulu archipelago.

The arbitrariness of population boundaries becomes more apparent with increasing evidence that humans have always moved about and intermixed. Previous studies reported that Filipino groups were descended from Austronesian-speaking farmers who mixed with established local hunter-gatherers in the Philippine archipelago (Lipson et al., 2014; Jinam et al., 2017). Quasi-racial categorizations akin to those applied to continental groups (Lewis et al., 2021) are even more inappropriate when applied to very localized geographic regions, as intermarriages and migrations between islands potentially blur any semblance of genetic divides. Given this background, it will not be surprising to find genetic evidence for the mixing of Sama and Tausug ancestors and other populations integrating into the melting pot that was the Sultanate of Sulu.

Extensive analysis of historical and anthropological literature and interactions with locals proved extremely helpful in understanding the contexts of group identities and ethnonyms. From the history of the last centuries, episodes of violence and political changes have influenced human migration in the Sulu Archipelago (Warren, 1985; Nimmo, 2001), and the sea has been an avenue of migration and interactions between islands. Notably, a few villages in Tawi-Tawi were founded by Tausug or Bangingi settlers from Sulu who mixed with the neighboring populace, adopted the Sama dialect of the vicinity, and currently identify as the local Sama. Tausug and Sama intermarriages are common, resulting in children who may identify as both. It is also not uncommon for an individual to have been born in a site different from where the field collection

was conducted or have one's grandparents originating from other islands.

Ethnonyms or group names in the Sulu archipelago are multi-layered, with a few eliciting pejorative connotations. For example, "Sama" is a broad term many participants identify with when asked about group affiliation. However, the Ethnologue distinguishes between Central Sama and Southern Sama languages, broadly corresponding to the geographic distribution of dialects spoken across Sulu and Tawi-Tawi (Eberhard et al., 2021). In the field, locals do not consciously adhere to linguistic classification but would instead use an island name (e.g., *Sama deya*, lit. Sama on land, *Sama bihing* lit. Sama onshore) to specify their group identity. Likewise, the Tausug (lit. People of Sulu) may distinguish between *Tau gimba* (forest people) or *Tau higad* (shore people) or refer to names of municipalities (e.g., *Tau Maimbung, Tau Parang*).

Sama Dilaut is an example of an ethnonym with various connotations. It is the name former boat dwellers would use to refer to themselves, meaning "Sama of the sea". The perceived eccentricity of "sea nomadism" practiced by the *Sama Dilaut*, their adherence to animistic beliefs, and their reluctance to integrate into mainstream society have contributed to their long history of social isolation and ostracism. Pejorative labels (e.g., *Luwaan* lit. spitted-out) were also sometimes applied to them (Nimmo, 2001). The *Sama Dilaut* is more commonly known by the exonym, Badjao (or Bajau), and many of them throughout the Philippines have adopted this name rather neutrally. However, researchers need to be aware of the various attitudes of other Sama (non-Sama Dilaut) communities towards being identified with the term. The attitudes range from its acceptance as being synonymous with Sama (especially in nearby Malaysia) to being appropriate only to sea nomads.

Moreover, a person's ethnic identity may sometimes have little to do with ancestry. In Mapun, many locals who descended from Tausug ancestors, having spent their entire lives on the island and primarily speaking Jama Mapun, identify as Jama Mapun. JJRBR and JMDC later learned that "Jama" means "person" and "Jama Mapun" means a person from Mapun.

Consequently, ethnic identities in this region are not as neatly grouped as would be convenient for data analysis. The multi-layered nature of terms and the fluidity by which individuals self-identify in Sama and Tausug communities suggest that ethnic labels are not perfectly congruent with biological affinities and must be viewed with the awareness of potential bias in using these terms to categorize data. Individuals with varying levels of admixture contributed samples to this dataset - from those whose grandparents reside on the same island, to individuals who describe themselves as broadly Sama in ancestry but whose ancestors originate from different islands, to those with multiple ethnicities in their family tree. Our approach is to explore ways data can be arbitrarily grouped and examine which sets more faithfully mirror genetic clusters. We further support the recommendation to use endonyms or self-applied group names in publications, or in cases when the complete consensus among the group is lacking, the use of neutral exonyms provided with explanatory notes (Max Planck Institute for Evolutionary Anthropology, 2021).

Upon returning to the field sites, the research team will work with local partners to increase the communities' understanding of what science can and cannot explain. Researchers will discuss the limitations of genomic inquiry and emphasize relatedness, the multiplicity of ancestors, and movements and intermixing of peoples in the past and present. As the spread of languages or cultural features may not be accompanied by the spread of genes (Diamond and Bellwood, 2003), it will be emphasized that results are not expected to fit perfectly with written or oral history, nor does genetic ancestry confer identity. The information campaign will be done with Indigenous groups, local partners in MSU-TCTO, NCIP, OSCC, MIPA, and the general public.

4.4 Diverse sectoral appreciation of the Indigenous peoples' ownership of their samples and associated data

The ownership of biological samples and the derived genetic information is an integral component of the informed consent of human participants in a research study. However, Indigenous peoples, many of whom are socially and economically vulnerable, are not familiar with the process and may have unknowingly provided broad consent for their samples to be used and stored in local and global databases. While data sharing, rapidly publishing results, and follow-up investigations are essential to advance the field, broad consent and unrestricted access to genomic data is counter to the Indigenous peoples' autonomy over their genomic data (Garrison et al., 2019) and excludes communities from sharing in potential benefits (Tsosie et al., 2021). Even with individual anonymization, social harms arising from the unregulated use of genomic data cannot be precluded, given that such data are usually linked to information about group affiliation.

To protect the rights of persons and their communities, the United Nations issued the Declaration on the Rights of Indigenous Peoples in 2007, which stated that "Indigenous peoples have the right to maintain, control, protect and develop their cultural heritage, traditional knowledge and traditional cultural expressions, … including human and genetic resources" (p. 3). In this framework, individuals and their communities own their genomic data, which must be recognized by the researchers conducting the primary study, and in all subsequent research arising from further use of archived biological samples and the derived genetic data. Considering secondary data usage, a system employing controlled access sharing (Byrd et al., 2020), where an access committee evaluates whether the requesting party substantially deviates from the terms of

the original consent (thus warranting new iterative consultations with the community), is desirable. Data access committees must include members from Indigenous communities or government agencies mandated to protect such groups. Local leaders or representatives should foster accountability from researchers and ensure that their ownership and autonomy over their genetic data are genuinely respected.

In this study, the individual and community ownership of samples and derived genetic data is recognized and documented in the MoA signed by the University of the Philippines, the NCIP, and the recognized leaders of each Indigenous group. Moreover, the protocol for handling, storing, and analyzing data was submitted to the UPMREB, which monitors the research team's compliance with the approved protocol. This level of transparency in the use of samples and genetic information within the parameters of the consent provided clearly manifests researchers' recognition of the Indigenous peoples' ownership of their samples and genetic information.

4.5 Varying compliance with ethical requirements by researchers and other stakeholders

Ethical review aims to balance the need to protect human participants from possible harm with the conduct of research that is beneficial to the community and the general public. Ethical issues have been raised concerning scientists from high-income nations conducting research among Indigenous peoples from low- and middle-income countries (LIMC) (van Teijlingen and Simkhada, 2012; Lahey, 2013; Pasic et al., 2018; Schroeder et al., 2018). All too often, this involves "helicopter research", i.e., research conducted under different ethical standards or with less oversight, with the source country having little say in the types of studies conducted, access to biological samples and study findings, and level of benefit sharing with the participants and their communities (Nature Journal, 2022).

As local or international scientists may choose not to follow regulations for protecting human participants, the responsibility for adhering to such rules must be shared by other stakeholders (De Ungria and Jimenez, 2022). For example, scientific journals are responsible for declining the publication of studies or retracting works where ethical misconduct has been demonstrated. Most journals uphold the protection of human participants as stipulated in the Declaration of Helsinki. However, this is primarily intended for physicians conducting patient research. As the scope of research involving humans extends to other disciplines beyond medical research, journals should consider upholding other relevant international ethical guidelines (De Ungria and Jimenez, 2021). In particular, the 2016 Declaration of Taipei (World Medical Association, 2016) and the CIOMS guidelines (Council for International Organizations of Medical Sciences, 2016) cover issues relating to biobanking, databases, and various aspects of human research. Journals should also require ethical compliance to globally accepted principles and legal guidelines in countries from which those genetic samples were sourced. In June 2022, Nature reported that it is improving its policies for publishing papers, following the recommendations made during the seventh World Conference on Research Integrity (Nature Journal, 2022). However, the scope of the policy review was unclear regarding the increased protection of the human participants, including Indigenous peoples, who in many LMICs are more socially and economically vulnerable than the general population.

Additionally, funding agencies are responsible for only supporting research that is ethical and investigating any allegations of misconduct by investigators of projects they finance. For example, the European Commission upholds the Global Code of Conduct for Research in Resource-Poor Settings (Schroeder et al., 2019) as a mandatory reference for the research and development activities it supports. Among other guidelines, the code admonishes researchers to seek ethical review in the host country, when available, even if an ethics approval has already been obtained in the researchers' home country.

Lastly, research institutions have administrative oversight over their affiliated scientists and should ensure that laws and regulations are observed correctly in jurisdictions where they conducted their field studies. Such institutions should proactively conduct fair and thorough investigations whenever misconduct accusations arise. Some recent studies on Philippine Indigenous communities allegedly did not fully comply with local policies and proceeded to conduct research without the proper clearance from institutions accredited by PHREB (Rochmyaningsih, 2022). The highest level of accountability for this work is expected from scientists who continue to claim that they have the required ethics and institutional clearances, even after being informed that this was not the case (National Commission on Indigenous Peoples, 2015; National Commission on Indigenous Peoples, 2021a; Philippine Genome Center, 2021).

While it is the primary responsibility of researchers to ensure that the conduct of research is ethical, a research's social and scientific value is best achieved through a multi-stakeholder cooperation. This includes research institutions, funding agencies, ethics review committees, government institutions, and scientific journals (De Ungria and Jimenez, 2022). This research is being conducted under the monitoring of the UPMREB, the NCIP, and OSCC/MIPA. Moreover, reports are submitted annually to the UPMREB to document our adherence to the approved protocol. This protocol underwent an ethics review before the start of the project. Any change or amendment to protocol also undergoes an ethics review process by the committee composed of social and natural scientists, lay people, and ethics experts before implementation to ensure the protection of the study participants.

During the fieldwork, research results were communicated in the local language. Forms and documents were translated,

and scientific concepts were explained through visual aids. It was nonetheless evident in many instances that the community did not fully comprehend all scientific technicalities. In this study, the participant's willingness to be part of a population genetics study relied heavily on the community's trust in the researchers and our partners in MSU-TCTO, OSCC, and NCIP. A response often received roughly translated to: "For as long as it is for the common good, then I am willing to take part." Valuing reciprocity in research means being duty-bound to maintain and honor the people's trust during and after the study for as long as genomic data remains accessible for analysis and interpretation. Because of these field experiences where the communities shared not only their samples but also their trust with the research team, we highly recommend that scientists, particularly those who engage with Indigenous peoples aspire to set an example of good research practice in the community. The research must be characterized by transparency, honesty, accountability, and regulatory compliance as concrete manifestations of how scientists should value the trust afforded by Indigenous persons and their communities.

4.6 Need to provide direct social benefits to the research participants and their communities

In living in the Sama and Tausug communities, the research team became familiar with their pressing needs, including necessities such as access to healthcare, education, and means of livelihood. Poor understanding of Indigenous cultures, even by fellow Filipinos, remains a drawback, resulting in the misrepresentation of Indigenous peoples in government and the public domain. This lack of representation often manifests in governance that fails to prevent further displacement and loss of cultural identity and ancestral domains (Seitz, 1998; Human Rights Watch, 2014; Chandran, 2018).

As academic researchers, it is well beyond our capacity to implement systemic changes that ensure these needs are met. The social aspect of population genetic research and the value of interacting with communities beyond the normal boundaries of scientific research are too often overlooked especially with the intensifying race among scientists to publish novel findings. However, scientists must remember that maximizing beneficence and reducing the risk of social harm are fundamental principles of ethical research (World Medical Association, 2013). We, therefore, recommend that researchers find creative ways of helping the communities by tapping resources, particularly those from the educational and health sectors, and communicating through their extensive networks to connect agencies and people that could assist Indigenous peoples.

In addition, education and building research capacity are areas where communities can draw much social benefit (Claw et al., 2018). In some countries, scientists foster partnerships with Indigenous peoples through scientific training and internship programs with the long-term goal of equipping Indigenous scientists to carry-out genomic research to benefit their communities (e.g., Carl R, 2018). Efforts of this type of engagement were made in partnership with the MSU-TCTO. The research team delivered lectures on genetics, linguistics, and archaeology to students and the teaching staff of MSU-TCTO, many of whom are members of the Sama and Tausug communities. In seeking their assistance, the research team aims to understand how the Indigenous peoples view the results of this study using their cultural lens to design better information materials for the communities. Moreover, the research team has committed to crafting a university-level Sama studies course as part of the project in line with the CHED's policies to integrate Indigenous people's studies into the higher education curricula (Commission on Higher Education, 2019).

5 Conclusion

Genomic technologies applied to studies of present-day and ancient populations offer unprecedented insights into our history as a human species. Whereas it could be tempting to view national ethical procedures as bureaucratic obstacles, we believe that prioritizing the welfare of communities increases the research's overall scientific and social value. While the field has rapidly advanced, Indigenous communities who are the focal subjects of such studies, remain vulnerable and marginalized, especially in low- and middle-income settings. To bridge this gap, it is crucial to recognize Indigenous peoples as partners in the research process, as rightful owners of genetic information, and place them on the receiving end of benefits derived from the study. Moreover, it cannot be overemphasized that this is a multi-stakeholder pursuit where academic institutions, government agencies, and journals play vital roles in ensuring that the highest ethical standards are upheld. Lastly, recognizing that every Indigenous group is unique, we encourage ethical, legal, and social implications studies to be conducted during community engagements. These would entail delving into local perceptions of communities regarding consent, traditional knowledge, ownership, and their needs and values, among others.

This paper aims to contribute knowledge to researchers, regulatory agencies, and policymakers by sharing our experiences and making recommendations to address our challenges during fieldwork. We call for revisiting existing regulations to enhance the protection of human participants, especially those who belong to Indigenous groups while creating an enabling research environment for the benefit of all stakeholders. Overall, these are essential steps toward improving the governance of ethical research in the Philippines and, possibly, in other countries undergoing similar

challenges, generating research outcomes that are meaningful and respectful of the rights and needs of Indigenous peoples.

Author contributions

JR conceptualized the study with input from MS and MD. Community consultations and sample/data collections were carried out by JR, JC, MH, LZ, and AI. RM and AI constitute the research team's counterpart in the Sulu Archipelago. JR, JC, EJ, MS, and MD reviewed and critically discussed pertinent rules and guidelines. All authors discussed and agreed with the recommendations in the final manuscript.

Funding

This work is financially supported by the Salikha Creative Grants Program under the Commission on Higher Education and the National Commission for Culture and the Arts (SALIKHA74UPD), the Department of Science and Technology – Philippine Council for Health Research and Development (DOST-PCHRD) through the Filipino Genomes Research Program (FGRP), the Natural Sciences Research Institute (NSRI-22-1-01), and the Max Planck Institute for Evolutionary Anthropology (MPI-EVA).

Acknowledgments

We thank Abdulaziz M. Joe and Madkie A. Biste of the Office for Southern Cultural Communities (OSCC) as well as Former Director Henrico L. Gumibao, Former Director Atty. Pinky Grace P. Pareja, OIC Director Mercilyn B. Hatad, Atty. Gibran Abubakar, and Engr. George T. Jocutan of the National Commission on Indigenous Peoples—Region IX. We also acknowledge our partnership with the Mindanao State University—Tawi-Tawi College of Oceanography and Technology (MSU-TCTO) through Chancellor Mary Joyce Guinto-Sali, Prof. Mucha-Shim L. Quiling, Prof. Abdul Jim J. Hassan, and Sitti Zayda B. Halun. Our fieldwork team was composed of Isnaji K. Nurlaji (OSCC-Tawi-Tawi), Faizal Dawamie (OSCC-Sulu), Farhanie Bernardo, Thelma Alman, Jhay Sakandal, Winnie Perian, Richard Pamaran, Czarina Mae M. Malun, Imelda M. Velasquez. Nick O. Acdalin (NCIP), Khalid D. Adam, Alimar J. Sakilan, (MSU-TCTO), Alpha Rae M. Espigar (UPD-ASP), Noriel B. Esteban, and Kathleen Crizelda M. Guillo (UPD-NSRI DAL FGRP). We also thank the Philippine Coast Guard, Southwestern Mindanao, Philippine Navy, and Philippine National Police Sulu. We also thank Jainor Timothy U. Garcia, Jeremie Jane A. Rivera, Joanne M. Santos, Ramil N. Palonson, Minerva Sagum, and our colleagues at the UPD-NSRI DAL and the Population History Group at the MPI-EVA. The authors acknowledge Jacob Barbosa for the illustration and Jacob Maentz for the photograph used in Figure 3. Lastly, we thank the people and local government units of Zamboanga and the Sulu Archipelago for their warm hospitality and generosity.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Barrows, D. P. (1910). The Negrito and allied types in the Philippines. *Am. Anthropol.* 12, 358–376. doi:10.1525/aa.1910.12.3.02a00020

B. Bellina, R. Blench, and J. C. Galipaud (Editors) (2021). *Sea nomads of Southeast Asia from the past to the present* (Singapore: NUS Press).

Bellwood, P. (1997). *Prehistory of the indo-Malaysian archipelago*. United States: Univ. Hawai'i Press.

Blanchard, J. W., Outram, S., Tallbull, G., and Royal, C. (2019). We don't need a swab in our mouth to prove who we are": Identity, resistance, and adaptation of genetic ancestry testing among native American communities. *Curr. Anthropol.* 60 (5), 637–655. doi:10.1086/705483

Blust, R. (2007). The linguistic position of Sama-Bajaw. *Stud. Philipp. Lang. Cult.* 15, 73–114.

Blust, R. (1995). The prehistory of the austronesian-speaking peoples: A view from language. *J. World Prehist.* 9, 453–510. doi:10.1007/bf02221119

Byrd, J. B., Greene, A. C., Prasad, D. V., Jiang, X., and Greene, C. S. (2020). Responsible, practical genomic data sharing that accelerates research. *Nat. Rev. Genet.* 21, 615–629. doi:10.1038/s41576-020-0257-5

Carl, R. (2018). Woese Institute for genomic BiologySummer internship for INdigenous peoples in genomics (SING). https://sing.igb.illinois.edu (Accessed January 9, 2022).

Chandran, R. (2018). Driven from home, Philippine Indigenous people long for their land. https://www.reuters.com/article/us-philippines-landrights-crime-idUSKBN1HQ034 (Accessed January 14, 2022).

Chennells, R., and Steenkamp, A. (2018). "International genomics research involving the san people," in *Ethics dumping case studies from north-south research collaborations. SpringerBriefs in research and innovation governance.* Editors D. Schroeder, J. Cook, F. Hirsch, S. Fenet, and V. Muthuswamy (Cham: Springer). doi:10.1007/978-3-319-64731-9_3

Claw, K. G., Anderson, M. Z., Begay, R. L., Tsosie, K. S., Fox, K., Garrison, N. A., et al. (2018). A framework for enhancing ethical genomic research with Indigenous communities. *Nat. Commun.* 9, 2957. doi:10.1038/s41467-018-05188-3

Commission on Higher Education (2019). CHED memorandum order no. 02 series of 2009. https://ched.gov.ph/wp-content/uploads/CMO-No.-02-Integration-of-Indigenous-peoples-studies-into-the-relevant-Higher-Education-Curricula.pdf (Accessed January 14, 2022).

Council for International Organizations of Medical Sciences (2016). International ethical guidelines for health-related research involving humans. https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf (Accessed January 14, 2022).

De Ungria, M. C. A., and Jimenez, E. B. (2022). "Ethical governance of forensic DNA databases in Southeast Asia," in *Handbook of DNA profiling*. (Gateway East, Singapore: Springer Nature Singapore Pte Ltd). doi:10.1007/978-981-15-9364-2_9-1

De Ungria, M. C. A., and Jimenez, E. B. (2021). Quezon City.Ethical challenges in genetics research with Indigenous cultural communities/Indigenous Peoples in the Philippines. Philippine Genome Center

Delfin, F., Salvador, J., Calacal, G., Perdigon, H. B., Tabbada, K. A., Villamor, L. P., et al. (2011). The Y-chromosome landscape of the Philippines: Extensive heterogeneity and varying genetic affinities of negrito and non-negrito groups. *Eur. J. Hum. Genet.* 19, 224–230. doi:10.1038/ejhg.2010.162

Delfin, F. S., Ko, A.-M.-S., Li, M., Gunnarsdóttir, E. D., Tabbada, K. A., Salvador, J. M., et al. (2014). Complete mtDNA genomes of Filipino ethnolinguistic groups: A melting pot of recent and ancient lineages in the asia-pacific region. *Eur. J. Hum. Genet.* 22 (2), 228–237. doi:10.1038/ejhg.2013.122

Department of Science and Technology (2012). Joint memorandum 2012-001 requirement for ethical review of health research involving human participants. https://ethics.healthresearch.ph/index.php/phoca-downloads/category/1-memorandum?download=1:Joint%20Memo%20Ethics (Accessed June 27, 2022).

Détroit, F., Mijares, A. S., Corny, J., Daver, G., Zanolli, C., Dizon, E., et al. (2019). A new species of *Homo* from the Late Pleistocene of the Philippines. *Nature* 568, 181–186. doi:10.1038/s41586-019-1067-9

Diamond, J., and Bellwood, P. (2003). Farmers and their languages: The first expansions. *Science* 300, 597–603. doi:10.1126/science.1078208

David M. Eberhard, Gary F. Simons, and Charles D. Fennig (Editors) (2021). *Ethnologue: Languages of the World*. Twenty-fourth edition (Dallas, Texas: SIL International).

Garrison, N. A., Barton, K. S., Porter, K. M., Mai, T., Burke, W., and Carroll, S. R. (2019). Access and management: Indigenous perspectives on genomic data sharing. *Ethn. Dis.* 29 (3), 659–668. doi:10.18865/ed.29.S3.659

GenomeAsia100K Consortium (2019). The GenomeAsia 100K Project enables genetic discoveries across Asia. *Nature* 576, 106–111. doi:10.1038/s41586-019-1793-z

Gray, R., Drummond, A., and Greenhill, S. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323, 479–483. doi:10.1126/science.1166858

Gunnarsdóttir, E. D., Li, M., Bauchet, M., Finstermeier, K., and Stoneking, M. (2011). High-throughput sequencing of complete human mtDNA genomes from the Philippines. *Genome Res.* 21, 1–11. doi:10.1101/gr.107615.110

Human Rights Watch (2014). Philippines: Protect zamboanga's displaced minorities. https://www.hrw.org/news/2014/04/30/philippines-protect-zamboangas-displaced-minorities (Accessed January 14, 2022).

International Alert (2020). Conflict alert 2020: Enduring wars. https://reliefweb.int/sites/reliefweb.int/files/resources/Philippines-Conflict-Alert-2020-Enduring-Wars-EN-2021.pdf (Accessed January 9, 2022).

Jinam, T. A., Phipps, M. E., Aghakhanian, F., Majumder, P. P., Datar, F., Stoneking, M., et al. (2017). Discerning the origins of the negritos, first sundaland people: Deep divergence and archaic admixture. *Genome Biol. Evol.* 9 (8), 2013–2022. doi:10.1093/gbe/evx118

Lahey, T. (2013). The ethics of clinical research in low- and middle-income countries. *Handb. Clin. Neurol.* 118, 301–313. doi:10.1016/B978-0-444-53501-6.00025-1

Larena, M., McKenna, J., Sanchez-Quinto, F., Bernhardsson, C., Ebeo, C., Reyes, R., et al. (2021b). Philippine Ayta possess the highest level of Denisovan ancestry in the world. *Curr. Biol.* 31 (19), 4219–4230.e10. e10. doi:10.1016/j.cub.2021.07.022

Larena, M., Sanchez-Quinto, F., Sjödin, P., McKenna, J., Ebeo, C., Reyes, R., et al. (2021a). Multiple migrations to the Philippines during the last 50, 000 years. *Proc. Natl. Acad. Sci. U. S. A.* 118 (13), e2026132118. doi:10.1073/pnas.2026132118

Lewis, A. C. F., Molina, S. J., Appelbaum, P. S., Dauda, B., Di Rienzo, A., and Fuentes, A. (2021). Getting genetic ancestry right for science and society. Available at: https://arxiv.org/abs/2110.05987 (Accessed January 4, 2022).

Lipson, M., Loh, P. R., Patterson, N., Moorjani, P., Ko, Y. C., Stoneking, M., et al. (2014). Reconstructing austronesian population history in Island Southeast Asia. *Nat. Commun.* 5, 4689. doi:10.1038/ncomms5689

Max Planck Institute for Evolutionary Anthropology (2021). Limitations of interpreting genetic data in light of human identity. https://www.eva.mpg.de/archaeogenetics/ethics/chapter-3 (Accessed January 9, 2022).

McKay, D. L. (2021). Real Indians: Policing or protecting authentic indigenous identity? *Sociol. Race Ethn. Thousand. Oaks.* 7 (1), 12–25. doi:10.1177/2332649218821450

National Commission on Indigenous Peoples (2021b). IP groups. https://ncip.gov.ph/ip-groups (Accessed January 9, 2022).

National Commission on Indigenous Peoples (2015). NCIP memorandum re: Genome research by larena 2015. https://ethics.healthresearch.ph/index.php/phoca-downloads/category/39-national-commission-on-indigenous-peoples-ncip?download=126:ncip-memo-re-genome-research-larena-2015 (Accessed January 9, 2022).

National Commission on Indigenous Peoples (1997). Republic act 8371 (indigenous peoples rights act). https://ncip.gov.ph/republic-act-8371/(Accessed June 27, 2022).

National Commission on Indigenous Peoples (2021a). Statement condemning the conduct of genetic/genomic research with Indigenous peoples by Dr. Maximillian Larena and his team without FPIC and the required ethical clearance. https://ncip.gov.ph/news/statement-condemning-the-conduct-of-genetic-genomic-research-with-indigenous-peoples-by-dr-maximillian-larena-and-his-team-without-fpic-and-the-required-ethical-clearance (Accessed January 9, 2022).

National Commission on Indigenous Peoples (2012a). The Indigenous knowledge systems and practices (IKSPs) and customary Laws (CLs) research and documentation guidelines of 2012. https://ncip.gov.ph/wp-content/uploads/2020/09/ncip-ao-no-1-s-2012-iksp.pdf (Accessed January 10, 2022).

National Commission on Indigenous Peoples (2012b). The revised guidelines on free and prior informed consent (FPIC) and related processes. https://ncip.gov.ph/wp-content/uploads/2020/09/ncip-ao-no-3-s-2012-fpic.pdf (Accessed January 10, 2022).

Nature Journal (2022). Nature addresses helicopter research and ethics dumping. *Nature* 606, 7. doi:10.1038/d41586-022-01423-6

Nimmo, H. A. (2001). *Magosaha: An ethnography of the tawi-tawi Sama Dilaut*. Quezon City: Ateneo De Manila University Press.

O'Connell, J. F., and Allen, J. (2004). Dating the colonization of Sahul (pleistocene Australia–new Guinea): A review of recent research. *J. Archaeol. Sci.* 31, 835–853. doi:10.1016/j.jas.2003.11.005

Official Gazette (2018). Republic Act No. 11054: An act providing for the organic law for the Bangsamoro autonomous region in Muslim Mindanao. https://www.officialgazette.gov.ph/downloads/2018/07jul/20180727-RA-11054-RRD.pdf (Accessed March 5, 2022).

Pagani, L., Lawson, D., Jagoda, E., Morseburg, A., Eriksson, A., Mitt, M., et al. (2016). Genomic analyses inform on migration events during the peopling of Eurasia. *Nature* 538, 238–242. doi:10.1038/nature19792

Pallesen, A. K. (1985). "Culture contact and language convergence," in *Linguistic society of the Philippines special monograph issue no. 24* (Manila: Linguistic Society of the Philippines).

Pasic, M. B., Vidrih, B., Sarac, H., Pasic, H., Vujevic, L., Soldo Koruga, A., et al. (2018). Clinical trials in developing countries - ethical considerations. *Psychiatr. Danub.* 30 (3), 285–291. doi:10.24869/psyd.2018.285

Philippine Genome Center (2021). Response to the June 9 2021 statement of Dr. Maximilian Larena and Prof. Mattias Jakobson on the conduct of their genomic study on the Filipino past. https://pgc.up.edu.ph/pgc-statement-july-23-2021-genomic-study-on-the-filipino-past (Accessed January 9, 2022).

Philippine Health Research Ethics Board (2017). National ethical guidelines for health and health-related research Manila: Philippine Council for health and research development. https://ethics.healthresearch.ph/index.php/phoca-downloads/category/4-neg (Accessed January 9, 2022).

Philippine Health Research Ethics Board (2016). PHREB-NCIP mou. https://ethics.healthresearch.ph/index.php/phoca-downloads/category/39-national-commission-on-indigenous-peoples-ncip?download=123:phreb-ncip-mou (Accessed January 9, 2022).

Philippine National Health Research System (2013). The implementing rules and regulations of republic act 10532, otherwise known as the "philippine national health research system Act of 2013. https://ethics.healthresearch.ph/index.php/phoca-downloads/category/17-pnhrs-law?download=103:pnhrs-irr (Accessed January 9, 2022).

Rochmyaningsih, D. (2022). Genomics' ethical gray areas are harming the developing world. https://undark.org/2022/02/24/genomics-ethical-gray-areas-are-harming-the-developing-world/?fbclid=IwAR3XospMJf7AguD6GCnoL7M9lnrtVdNxFQE495iRdLZWfYbj6-5B9EhD79Q (Accessed March 5, 2022).

Rohde, D. L. T., Olson, S., and Chang, J. T. (2004). Modelling the recent common ancestry of all living humans. *Nature* 431, 562–566. doi:10.1038/nature02842

Sather, C. (1997). *The Bajau Laut: Adaptation, history, and fate in a maritime fishing society of south-eastern Sabah*. New York: Oxford University Press.

Schroeder, D., Chatfield, K., Singh, M., Chennells, R., and Herissone-Kelly, P. (2019). "A value-based global code of conduct to counter ethics dumping," in *Equitable research partnerships. SpringerBriefs in research and innovation governance*. Editors D. Schroeder, K. Chatfield, M. Singh, R. Chennells, and P. Herissone-Kelly (Cham: Springer). doi:10.1007/978-3-030-15745-6_2

Schroeder, D., Cook, J., Hirsch, F., Fenet, S., and Muthuswamy, V. (2018). *SpringerBriefs in research and innovation governance*. Cham: Springer. doi:10.1007/978-3-319-64731-9Ethics dumping case studies from north-south research collaborations

Seitz, S. (1998). Coping strategies in an ethnic minority group: The aeta of mount pinatubo. *Disasters* 22 (1), 76–90. doi:10.1111/1467-7717.00076

Skoglund, P., Posth, C., Sirak, K., Spriggs, M., Valentin, F., Bedford, S., et al. (2016). Genomic insights into the peopling of the southwest pacific. *Nature* 538, 510–513. doi:10.1038/nature19844

Sterling, R. L. (2011). Genetic research among the Havasupai–a cautionary tale. *Virtual Mentor* 13 (2), 113–117. doi:10.1001/virtualmentor.2011.13.2.hlaw1-1102

Tabbada, K. A., Trejaut, J., Loo, J. H., Chen, Y. M., Lin, M., Mirazon-Lahr, M., et al. (2010). Philippine mitochondrial DNA diversity: A populated viaduct between taiwan and Indonesia? *Mol. Biol. Evol.* 27 (1), 21–31. doi:10.1093/molbev/msp215

TallBear, K. (2007). Narratives of race and indigeneity in the genographic project. *J. Law Med. Ethics* 35, 412–424. doi:10.1111/j.1748-720X.2007.00164.x

Tan, M. L. (2008). *Revisiting usog, pasma, kulam*. Quezon City: University of the Philippines Press.

Tsosie, K. S., Fox, K., and Yracheta, J. M. (2021). Genomics data: The broken promise is to indigenous people. *Nature* 91 (7851), 529. doi:10.1038/d41586-021-00758-w

United Nations (2007). United nations declaration on the rights of indigenous peoples. https://undocs.org/A/RES/61/295 (Accessed January 10, 2022).

van Teijlingen, E. R., and Simkhada, P. P. (2012). Ethical approval in developing countries is not optional. *J. Med. Ethics* 38 (7), 428–430. doi:10.1136/medethics-2011-100123

Warren, J. F. (1985). *The Sulu zone 1768-1895*. Quezon City: New Day Publishers.

World Health Organization (2006). The constitution of the world health organization. https://www.who.int/governance/eb/who_constitution_en.pdf (Accessed January 9, 2022).

World Medical Association (2016). Declaration of Taipei on ethical considerations regarding health databases and biobanks. https://www.wma.net/policies-post/wma-declaration-of-taipei-on-ethical-considerations-regarding-health-databases-and-biobanks/(Accessed January 14, 2022).

World Medical Association (2013). World medical association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA* 310 (20), 2191–2194. doi:10.1001/jama.2013.281053

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 08 June 2023
DOI 10.3389/fgene.2023.843209

OPEN ACCESS

EDITED BY
Arslan A. Zaidi,
University of Pennsylvania, United States

REVIEWED BY
LaKisha David,
University of Pennsylvania, United States
Cesar Fortes-Lima,
Uppsala University, Sweden

*CORRESPONDENCE
Fatimah Jackson,
fatimah.jackson@howard.edu

RECEIVED 25 December 2021
ACCEPTED 18 April 2023
PUBLISHED 08 June 2023

CITATION
Jackson F, Clinton C and Caldwell J (2023), Core issues, case studies, and the need for expanded Legacy African American genomics.
Front. Genet. 14:843209.
doi: 10.3389/fgene.2023.843209

Core issues, case studies, and the need for expanded Legacy African American genomics

Fatimah Jackson[1]*, Carter Clinton[2] and Jennifer Caldwell[3]

[1]Department of Biology, Howard University, Washington, DC, United States, [2]Department of Biology, North Carolina State University, Raleigh, NC, United States, [3]Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, LA, United States

Introduction: Genomic studies of Legacy African Americans have a tangled and convoluted history in western science. In this review paper, core issues affecting African American genomic studies are addressed and two case studies, the New York African Burial Ground and the Gullah Geechee peoples, are presented to highlight the current status of genomic research among Africa Americans.

Methods: To investigate our target population's core issues, a metadatabase derived from 22 publicly accessible databases were reviewed, evaluated, and synthesized to identify the core bioethical issues prevalent during the centuries of the African American presence in North America. The sequence of metadatabase development included 5 steps: identification of information, record screening and retention of topic relevant information, identification of eligibility via synthesis for concept identifications, and inclusion of studies used for conceptual summaries and studies used for genetic and genomic summaries. To these data we added our emic perspectives and specific insights from our case studies.

Results: Overall, there is a paucity of existing research on underrepresent African American genomic diversity. In every category of genomic testing (i.e., diagnostic, clinical predictive, pharmacogenomic, direct-to-consumer, and tumor testing), African Americans are disproportionately underrepresented compared to European Americans. The first of our case studies is from the New York African Burial Ground Project where genomic studies of grave soil derived aDNA yields insights into the causes of death of 17th and 18th Century African Americans. In the second of our case studies, research among the Gullah Geechee people of the Carolina Lowcountry reveals a connection between genomic studies and health disparities.

Discussion: African Americans have historically borne the brunt of the earliest biomedical studies used to generate and refine primitive concepts in genetics. As exploited victims these investigations, African American men, women, and children were subjected to an ethics-free western science. Now that bioethical safeguards have been added, underrepresented and marginalized people who were once the convenient targets of western science, are now excluded from its health-related benefits. Recommendations to enhance the inclusion of African Americans in global genomic databases and clinical trials should include the following: emphasis on the connection of inclusion to advances in precision medicine, emphasis on the relevance of inclusion to fundamental questions in human evolutionary biology, emphasis on the historical relevance of inclusion for Legacy African Americans, emphasis on the ability of inclusion to foster expanded scientific expertise in the target population, ethical engagement with their

descendants, and increase the number of science researchers from these communities.

KEYWORDS

bioethics, NYABG, Gullah Geechee, genomic equity, genetic databases

Introduction: status of genomic studies among peoples of African descent

The representation and treatment of African Americans in the biomedical, anthropological, and genomic literature has a tumultuous history. This systematic review discusses the importance of increasing genomic research performed on and for populations from underrepresented ancestries and the significance this enrichment would bring to our global genomic databases and accelerate the progress of our science. If all populations were fairly represented, the potential benefits of genomic research (e.g., better understanding of disease etiology, earlier detection ad diagnostics, rational drug design, and increased clinical care) would be available to underrepresented groups such as peoples of African descent (Fatumo et al., 2022). However, the current persistent Eurocentric bias in genomics and genetics extends through all levels of the discipline, including affecting the utility of polygenic risk scores in disease studies (Martin et al., 2019), which still have limitations for peoples of African descent since such studies are rooted in GWAS databases with a North Atlantic ancestry-centered ascertainment bias (Gultig, 2023).

In this review, we focus on Legacy African American populations, i.e., indigenous African Americans living in the United States, as a subset of peoples of recent African descent, and the reoccurring deficit of meaningful and adequate genomic studies available on this group. Legacy African Americans are the historic African American descendants of the heritage of American Slavery, Jim Crow segregation, and institutionalized racial discrimination. African American genomics is best framed within the context of continental African genomic diversity. In this review, we explore the current core issues in African American genomic studies, using two specific case studies. Here we show that the addition of genomic data from the remains of deceased individuals is a valuable and necessary adjunct to those data derived from the biological samples of living individuals and that the aDNA provides more insights than simply relying on skeletal and dental assessments alone. Using the historic New York African Burial Ground (NYABG) and the contemporary Gullah Geechee peoples of the Coastal Sea Islands and South Carolina and Georgia Lowcountry, we demonstrate the potential benefits of collecting and exploring ancient DNA (aDNA) and modern DNA samples to create a robust database capable of stimulating future research on Legacy African Americans and beginning to bring parity to genomic inquiries.

Who are Legacy African Americans? According to historical sources (Eltis and Richardson, 2015; Eltis et al., 2017), their deepest ancestral origins go back to diverse regions of continental Africa. From 1,501 to 1867, enslaved Africans were forcibly and brutally embarked mainly from eight coastal regions of Africa. According to these historical records, 5.7% embarked from Senegambia, 3.2% from Sierra Leone, 2.7% from the Windward Coast, 9.6% from the Gold Coast, 16.1% from the Bight of Benin, 12.3% from the Bight of Biafra, 46.3% from West Central Africa, and 4.1% from Southeast Africa. Prior to the aggregation of kidnapped Africans at these export sites, the Africans were part of local empires and kingdoms throughout the continent. However, the West Central African coast was the largest slaving and embarkation region throughout most of the trans-Atlantic trade in enslaved Africans (Fortes-Lima and Verdu, 2021), and it focused mainly on groups living south of the Congo River. The Gold Coast, the Bight of Benin, and the Bight of Biafra became increasingly prominent slave collecting and embarkation regions after the mid-17th century, as the trans-Atlantic trade in enslaved Africans expanded with the growth of the Plantation Economy in the Americas.

In two recent articles (Caldwell and Jackson, 2021; Jackson, 2021), Legacy African Americans are identified as the current 40+ million Black Americans with multigenerational backgrounds (legacies) of extensive contact with the North American social, cultural, economic, and legal environments. During these 400 years of exposure or approximately 16 generations (25 years per generation) of direct contact with American Slavery, Jim Crow racism and segregation, disparate health and educational opportunities, this background has uniquely shaped both their genomes and epigenomes (Jackson et al., 2018a; Jackson et al., 2018b). African Americans are the third largest ethnic group in the United States and are the results of various admixture events, with today showing common ancestry with Africans (~82.1%), Europeans (~16.7%), and Native Americans (~1.2%) (Baharian et al., 2016).

Recent 20th century migrations, like the 1st and 2nd Great Migrations of African Americans, initiated greater intra-group genetic homogeneity despite populations being initially geospatially distant. For example, Detroit, MI attracted African American migrants from Louisiana and the Mississippi Delta. Chicago, IL disproportionately attracted Legacy African Americans from five counties in Mississippi. Los Angeles, CA attracted African Americans from east Texas and Louisiana with some stopping to found previously predominantly African American towns such Dearfield, CO, Nicodemus, KA, and McNary, AZ. Predominantly African American town are part of the history of America. Today, only thirteen historical African American towns survived, but their legacy of economic and political freedom is well remembered. The Oklahoma towns of Boley, Brooksville, Clearview, Grayson, Langston, Lima, Red Bird, Rentiesville, Summit, Taft, Tatums, Tullahassee, and Vernon, for example, attest to the historic settlement patterns of Legacy African Americans. African American-founded towns remained predominantly African American demographically until the towns were disbanded. Increasing urbanization of Legacy African Americans facilitated gene flow between microethnic

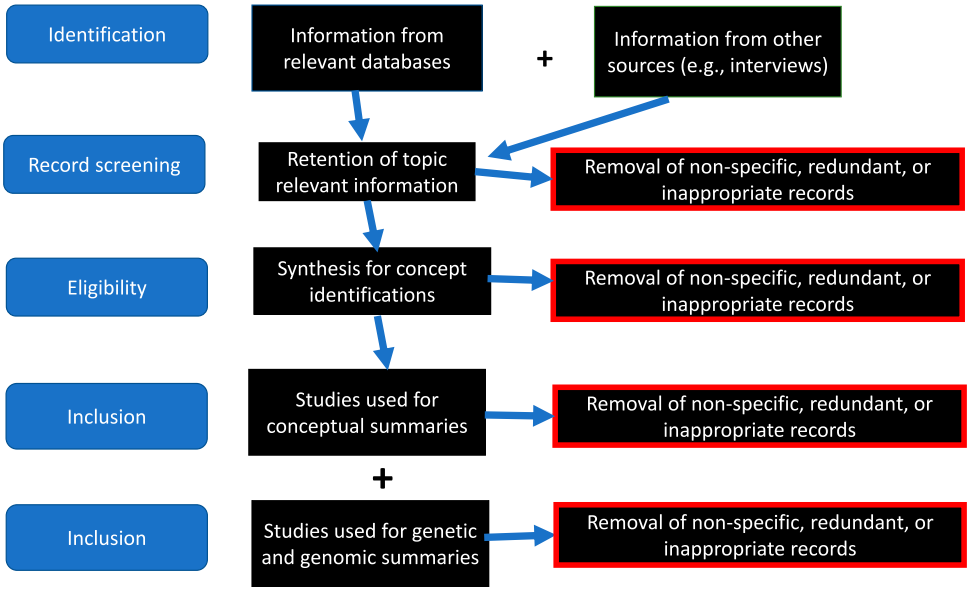


FIGURE 1
Sequence of development of metadatabase used in this study. All datatbases were publicly available; our strategy identified the most relevant information, refined and curated it, subjected it to uniformity procedures, and excluded al irrelevant data.

groups that had been previously distinct local communities with their own unique biological histories, subsistence patterns, and distinct African cultural retentions from the period of enslavement and its aftermath (*see* (Berlin, 2010)).

Methods and materials

As a systematic review, we accessed a wide variety of databases in this study to develop a megadatabase. The components of this mega database include.

- TransAtlantic Slave Voyages Database. (https://www.slavevoyages.org/voyage/database) (Slave Voyages Database).
- Accessible Archives Database https://www.accessible.com/accessible/preLog (Accessible Archives)
- AnthroSource, (https://anthrosource.onlinelibrary.wiley.com) (AnthroSource Database),
- Applied Social Sciences Index and Abstracts (ASSIA), (https://about.proquest.com/en/products-services/ASSIA-Applied-Social-Sciences-Index-and-Abstracts/) (Sociological Abstracts Database),
- socINDEX,https://proxy.library.emory.edu/login?url=http://search.ebscohost.com/login.aspx?authtype=ip,uid&profile=ehost&defaultdb=sih) (Emory university)
- Social Sciences Full Text (Wilson Web), (https://www.library.nd.edu/database/4qBOLHGnSwwM0ekI4SUkKC) (Hesburgh libraries),
- Google Scholar, (https://scholar.google.com/) (Google Scholar Search Engine),
- GenBank, (https://www.ncbi.nlm.nih.gov/genbank/) (GenBank),
- 1000 Genomes Project, (https://en.wikipedia.org/wiki/1000_Genomes_Project) (1000 Genomes Project),
- Online Mendelian Inheritance in Man (OMIM), (https://www.omim.org/) (OMIM),
- NCIB Reference Sequence (RefSeq) (https://www.ncbi.nlm.nih.gov/refseq/) (NCIB Reference Sequence)
- The World Bank Database (WHO), https://databank.worldbank.org/databases/africa) (The World Bank Database)
- Open Data for Africa, (https://dataportal.opendataforafrica.org/) (Open Data for Africa),
- African American Biographical Database, (https://aabd.chadwyck.com/) (The African American Biographical Database),
- African American Home Movie Archive, (https://www.aahma.org/) (African American Home Movie Archive),
- African American Odyssey, (https://memory.loc.gov/ammem/aaohtml/aohome.html) (Hine et al., 2010),
- Afro-American Genealogical Research: Introduction (https://guides.loc.gov/african-american-genealogical-research) (Afro-American Genealogical Research, 2020)
- Records of the Continental and Confederation Congresses and the Constitutional Convention (https://www.archives.gov/research/guide-fed-records/groups/360.html) (National Archivesa)
- Annals of Congress, Vol. 1: 1st through 18th (https://memory.loc.gov/ammem/amlaw/lwaclink.html) (The Library of Congress)
- RG 233: Records of the United States House of Representatives (https://www.archives.gov/research/guide-fed-records/groups/233.html) (National Archivesb)
- American Freedmen's Inquiry Commission, (http://www.shfg.org/resources/Documents/7-Strickland.pdf) (American Freedmen's Inquiry Commission), and

- Sociological Abstracts, (https://proquest.libguides.com/socabs) (Sociological Abstracts).

Data from these databases were explored and integrated into our discussion of an overall set of concepts and two specific case studies of the bioethics of African American genetics. We synthesized data from these sources as well as our emic perspectives to identify the core issues prevalent during the centuries of the African American presence in North America and germane to the bioethics of genomic biomedical research in this population. The sequence of our uses of these online databases are depicted in Figure 1. Our coordinated review of the constituent databases and other sources of relevance included five steps: identification of information, screening and data-transformation of the records, determination of underlying African American conceptual issues on genetics and genomics, and incorporation of existing genetic and genomic data on Legacy African Americans. Excluded from our consideration were studies on non-Legacy African Americans, studies that did not consider continental Africans, and studies that did not test specific genetic or genomic hypotheses.

Case studies in African American genomics

Case #1: New York African Burial Ground (NYABG)

The NYABG is the country's oldest and largest burial site of free and enslaved Africans ever discovered. Its origin dates to ~1,640, with a closing date around 1797. The site spans 6.6 acres across the New Amsterdam Colony or present-day New York City (NYC) (GSA). The initial use of the burial ground coincides with the establishment of the Negro Frontier, a free African community just outside of the New Amsterdam Colony. A community that needed a place nearby to bury their loved ones without having to carry them beyond the Colony's boundary walls or paying a fee. The NYABG was rediscovered in 1991, when 419 skeletons were unearthed during the construction of a federal building at 290 Broadwayin Lower Manhattan. Researchers at Howard University performed robust analyses and generated initial reports on the skeletal biology, history, and archaeology of the site and the population. At the conclusion of this landmark project, the skeletal remains were reburied out of respect leaving only the grave soil (collected simultaneously with the remains) for future study.

In 2015, we initiated a study investigating the soil chemistry and bacterial community diversity (including infectious disease pathogens) of the burial soil samples and their geospatial patterns. We have successfully detected all human-associated bacteria for each burial inhabitant. We have reconstructed the human microbiome for 66 NYABG individuals. Detection of human microbiome profiles gives us insight into individual and ancestral identity, living conditions, and possible causes of death of the corresponding burial inhabitant. Our findings demonstrate the capability to detect human evidence in soil that has been buried for 400-hundred-years. This demonstration serves as proof of concept to explore genomic human aDNA in the NYABG soil samples and other burial soils around the country of similar age (Clinton, 2021). Researchers have acknowledged the human microbiome as our "second genome," i.e., an additional source of genetic diversity and identity (Grice and Segre, 2012). The potential aDNA analysis from the NYABG soil samples will allow us to capture a subset of a historical population (15,000 still buried) and enrich genomic databases with African descended genomic data. By capturing the genomic architecture of this 17th and 18th population, we can perform population genetics analyses to observe evidence of human variation and disease susceptibilities helping us to combat health disparities. We can also contribute this newly generated data to existing databases where people of African descent are underrepresented. We hope to use this investigation as a proxy for the potential to explore other African American burial grounds around the country without disturbing or destroying remains but still learn all that we can about the genetics of African Americans.

Challenges of studying historic African American remains

Several challenges must be addressed when studying historical remains (skeletal or soil) of African American populations for ethical scientific research with advanced molecular technologies. One challenge is to ensure that research on historic populations is performed in an ethical nature by protecting the sacred ground where they are buried. In addition, legislation must be established to ensure construction or housing development projects do not decimate African American burials (Clinton and Jackson, 2021). Often, these projects physically destroy burial sites and erase the existence and contributions of the buried population from history. The lack of burial site protection for African Americans promulgates the idea that they, as a population (alive or dead), are worth less than other populations in America. The lack of protection increases the difficulty for researchers to gain access to these grounds and move forward with investigations. Another challenge of studying historical remains is determining the best research team to engage with the underrepresented community and conduct the research. Performative and helicopter research are two methods of conducting predatory research where the interpretations and generation of data do not benefit the studied population. These types of research promote more harm than good, resulting in the perpetuation of mistrust between marginalized communities and scientific researchers. The appropriate research team for studying African Americans and other underrepresented and marginalized communities are those who perform research for the greater good of the community, serving their needs for increased representation in databases, accurate interpretation of generated data, and ethical applications of the research to better health outcomes. It is paramount to consider the appropriate decision-makers for how to conduct the research. Decisions should be made by educated members of the descendant community, those who are likely closely genetically and culturally related to the studied population. In some cases, the descendant community may not be the local community but genetically related descendants some distance away from the site. A third challenge is establishing where the generated data will be housed and who can access it. We propose that researchers store

data in a private repository where access can be controlled by the stakeholders, i.e., the local or descendant population who will directly benefit from the research. Researchers must acknowledge the purpose, potential impacts, and sources for conducting research, generating data and reporting new findings with the scientific community, general public, and community upon which the research is performed (AABA Code of Ethics).

CASE #2: The Gullah Geechee population of the Carolina Lowcountry

The Gullah Geechee people are an historically important Legacy African American microethnic group residing largely in the Southeastern United States They are a candidate ancestral group to a diverse array of African American peoples across North America. The original migrations of the ancestral Africans moved from staging areas like Charleston, SC and the nearby Sea Island to more inward locales as the United States Frontier was pushed westward. In addition to their geographical isolation (Matory, 2008), enhanced retention of African allelic variants and cultural practices, many Gullah Geechee peoples migrated from coastal Carolinas to adjacent regions. Some Gullah Geechee who escaped enslavement, fled to join Black Seminole populations in the Spanish held territory of Florida. Creek Freedman, many derived from Gullah Geechee lineages became refugees on the Trail of Tears to Indian Territory (present day Oklahoma). When the United States government forced First Nations peoples to accept individual land allotments, many Freedmen established predominantly African American towns with other former enslaved African Americans of the Five Tribes. Here they settled together for mutual protection and economic security (Oklahoma Historical Society). Black Seminole Freedmen populations also founded communities in northern Mexico, Texas, Oklahoma, and Red Bays Settlement in Andros Island, Bahamas (Holm, 1983), escaping Florida after the first Seminole War. The notion of ancestral linkage or the Gullah Connection is commonly affirmed via historical and anthropological records, but little work has been done to confirm this genetically (Paris, 1995; Opala, 2009). We have hypothesized that Gullah Geechee genomic and cultural signals proliferate beyond their current geographical territories in diverse African American communities throughout North America (Caldwell and Jackson, 2021).

The Gullah Geechee homelands of the Lowcountry were the most affluent area of British North America during the colonial period and became an optimal site for African-derived cultures to thrive and adapt. The harsh subtropical climates, malaria transmitting mosquitoes, thick marsh and swamp lands would provide environmental insulation for the amalgamation of Africans to retain and synthesize their own cultural preferences. It also provided an ecological setting to which many Gullah Geechee peoples were preadapted genetically. Intentional admixture was encouraged by Europeans and European Americans to prevent slave revolts among newly arriving enslaved Africans (Scharff et al., 2010). However, many South Carolina slave owners would visit their plantations only as needed to avoid the harsh climates while maintaining the authority needed to ensure profitable production schedules. Physical and social isolation allowed for the unique Gullah Geechee culture to emerge as a synthesis of many African (and non-African) traditions. Their storytelling, veneration of the ancestors, belief in a higher power, and unique cultural attributes were all amplified in the setting of the Sea Islands. The emerging Gullah Geechee peoples, like their creole dialect, represents a unique fusion of Niger-Kordofan and Afro-Asiatic, Indo-European, and Southeast First Nations patterns.

Like their cultural retentions, isolation enhanced the potential for genetic drift in the population. Contemporary lineages from Sapelo Island may reflect the disproportionate influence of a founder, Bilal who became known over time as Bailey among the local Gullah Geechee peoples (Bailey, 1995). As a progenitor Legacy African American population, the Gullah Geechee should have retained ancestral markers with a stronger West and Central African signal compared to other Legacy African American populations whose African signals may have been diluted by more admixture with non-Africans. We suggest that the Gullah-Geechee genomic profiles will show distinct characteristics of endogamy and substructure when compared to other African American microethnic groups as a reflection of their unique history and preeminence. In a recent unrelated study of the Gullah-Geechee (Zimmerman et al., 2021) it was observed that, relative to non-Gullah African Americans from the Southeast United States, the Gullah exhibited higher mean African ancestry, lower European admixture, a similarly small Native American contribution, and increased male-biased European admixture. A slightly tighter bottleneck in the Gullah 13 generations ago suggests a largely shared demographic history with non-Gullah African Americans, as we observed previously (Caldwell and Jackson, 2021). Despite a slightly higher relatedness to populations from Sierra Leone (Zimmerman et al., 2021), overall, the studies demonstrate that the Gullah are genetically related to many African populations, representing an amalgamation of West and Central Africans in particular (Caldwell and Jackson, 2021).

A recent study (Zimmerman et al., 2021) confirms that subtle differences in African American population structure exist at finer regional levels. Such observations were reported decades ago (Jackson, 2004; Jackson, 2008)and their validation can help to inform medical genetics research in African Americans and guide the interpretation of genetic data used by African Americans seeking to explore ancestral identities.

Using the Ely-Jackson database, Bert Ely and others (Ely et al., 2006) completed a mtDNA analysis of 78 Legacy African Americans who lived in the Lowcountry and were considered Gullah Geechee descendants. 40% of participant Gullah Geechee had mtDNA migration patterns from West Central Africa, a proportion that resonates with our earlier studies of these peoples (Jackson, 2004; Jackson, 2008). Other Gullah Geechee mtDNA patterns were 23% from Senegambia and 18% from Upper Guinea. Over 30 percent of Ely's Gullah Geechee participants did not have a mtDNA match with their extensive database of over 4,000 African mtDNA variants (the Ely-Jackson Database), but were clearly of African origin (i.e., most were part of the L megahaplogroup). This illustrates the current limitations of the African-centered reference databases needed for comparative reconstructions of African origins. Although half of the African American participants were able to trace their ancestry to multiple ethnic groups of continental Africa south of the Sahara Desert, Ely and his team (Ely et al., 2006)

recognized that autosomal DNA would be needed to determine more information about the probably African ethnic groups of origin because mtDNA was not conclusive enough to determine a single ethnic source of maternal lineage. Ultimately, they suggested that more work should be done to geospatially map African American mtDNA haplotypes. In a more recent mtDNA study of the Gullah Geechee (Fleskes et al., 2021) all had mitochondrial lineages belonging to African haplogroups (L0-L3), with two individuals sharing the same non-African H1cb1a haplotype, while one had a Native American A2 mtDNA.

The geographic isolation of the Gullah Geechee well into the 21st century has allowed them to retain more African ancestry informative alleles and maintain more African cultural retentions than adjacent contemporary Legacy African Americans further inland. Our research among the Gullah Geechee has created a comprehensive analysis of this microethnic group to better understand how they evolved and impacted the broader African American communities. The genomic variance among the Gullah Geechee undoubtedly contribute to the dramatic patterns of health inequities in their region. Remnants of state-sponsored chattel slavery and draconian segregation laws relegated a large proportion of the African Americans to populate the Southeastern part of the United States densely and disproportionately. It is within these settings that various African American microethnic groups emerged and proliferated (Taylor, 2019). Generations later, descendants of enslaved Africans are still clustered in the Southeastern states (e.g., the Stroke Belt: North/South Carolina, Georgia, Florida, Arkansas, Louisiana, Mississippi Alabama). In these states, the prevalence of stroke, diabetes, and cardiovascular disease (CVD) are the highest nationwide (Barker et al., 2011; Karp et al., 2016). Advances in the control of modifiable biocultural risk factors that served as disease triggers given the genetic backgrounds and comorbidities of the Gullah Geechee (e.g., obesity, hypertension, cigarette smoking, and high salt diets), have produced a decline in stroke related mortality and morbidities. However, data continues to suggest major ethnic disparities in stroke related mortalities among African Americans. Heart disease is the number one killer of African American women (Esenwa et al., 2018). Poor CVD health in these communities is exasperated by the history of slavery, social segregation, lack of access to healthcare and healthcare providers, institutionalized racial discrimination, stress, and economic instability. Moreover, institutional levels of inequity, coupled with genomic mediators like the epigenome lead to physiological precursors for stroke (Kramer et al., 2017; Esenwa et al., 2018).

The scientific literature suggests that the high rates of chronic disease in African Americans are caused by the combined and compounded effects of genetic, environmental, and social factors. Yet little is known about the magnitude or geographical distribution of African American genetic diversity, cultural disease catalysts, and population substructure between and within African American populations. Due to their underrepresentation and the present bias toward European and European American genomics, research is needed to understand the effect of multiple genes, epigenetic modifiers, environment, and lifestyle and cultural risk factors that increase susceptibility of these multifactorial disorders (Grundy, 1998). We need to be able to apply network analysis and sophisticated computational biology models to depict interactions in African American populations.

In addition to considering the multiple contributing factors that influence chronic disease, our research among the Gullah Geechee suggests that ancestral analysis may uncover evolutionary contributions that have not been considered in other populations because of the ancient, frequently unacknowledged, and often unique genetic underpinnings of populations of recent African descent. Genome-wide studies (GWAS) have become important genomic tools to use in genetics to associate specific genetic variations with diseases. The method involves scanning the genomes from many different people and looking for genetic markers that can be used to predict the presence of a disease phenotype. Most GWAS are focused on Europeans (52%) and Asians (21%) (Hoffman et al., 2016a). African populations make up less than 1% of the total GWAS studies. The largest African American GWAS study consists of 8,000 individuals while the largest European American GWAS study encompasses 100,000 individuals (Abel and Schroeder, 2020). This means that means that many population-specific pathogenic variants are left undetected. Just as often, many alleles that could provide ameliorative effects for disease phenotypes also remain undiscovered. For example, African Americans are three times more likely to experience kidney failure than European Americans and African American kidney disease tracts clearly with dementia in African Americans (Laster et al., 2018; McAdams-DeMarco et al., 2018). Without the knowledge of the range of genomic diversity in our entire species, and particularly those individuals of recent African ancestry, our efforts to understand human variability adversely affects the control of associated health disparities, exaggerating these disparities over time, limiting the reproducibility of our data, and truncating the significance of our GWAS findings. Including more African-descended populations in genomic research widens the possibilities for more precise clinical application, biomedical treatments, evolutionary insights, and more equitable health policies for every population. Systematically Including African-descended groups takes the scientific community a giant step toward greater parity. For example, recent large studies (Tang et al., 2001) of GWAS for Alzheimer's Disease in African Americans found eleven novel risk loci, seven of which were rare. Many of the exact genes differed from those identified in European American GWAS investigations. This emphasizes the importance of using genomic studies to assess the higher dementia rates among African Americans and it confirms that the most important genes associated with Alzheimer's Disease vary between populations even though the deep ancestries of every human population can ultimately be traced to continental Africa.

Finally, our research among the Gullah Geechee suggests that an important avenue for exploring genomic and cultural variation in a geospatially complex and diffuse population such as Legacy African Americans is to study the founding population segments. The Gullah Geechee are an important African American founding population who emerged soon after Africans first were brought to the Carolina Lowcountry (Caldwell and Jackson, 2021). Researching such groups can provide important and unexpected insights into disease etiology and inheritance patterns. Two recent studies (Gupta, 2021; Zimmerman et al., 2021)confirms that subtle differences in African American population structure exist at finer regional levels, using the Gullah Geechee as an example, confirming the initial observations of substructure in African Americans in the

FIGURE 2
Blue highlighted areas are the geospatial locations of the two case studies reported in this paper. The New York African Burial Ground is located in New York City (Lower Manhattan) while the Gullah Geechee peoples reside along the Carolina Lowlands from Wilmington, NC to Jacksonville, FL.

United States. Such observations can help to inform medical genetics research in African Americans and guide the interpretation of genetic data used by African Americans seeking to explore ancestral identities. The genomics of founder populations can provide explanations for variations seen in complex disease mapping. Such efforts can also track the effects of genetic drift events and historical processes on the population, document regional changes in allele frequencies, identify evidence of

cultural adaptations, and monitor the incidence and prevalence of complex disease distributions. Founding events can also be used to locate progenitor populations for contemporary admixed populations. Investigation of founding populations and more inclusive GWAS studies have the potential to capture a wide range of genetic and environmental interaction networks while appropriately contrasting estimates of genetic risk versus environmental or systematic infrastructural risks that perpetuate current disadvantageous outcomes.

Figure 2 depicts the geographical ranges of the two case studies presented in this paper.

Core issues emerging from the case studies

Recently, the American Anthropological Association admitted to the racist attitudes and perceptions permeating the discipline with respect to the indigenous peoples of the Americas. As Gupta (Jackson, 1998) writes "*Since its inception, the history of American anthropology has been intertwined with a record of extractive research conducted on the Indigenous communities. Anthropologists have often assigned themselves the status of 'expert' over the cultural narratives and social histories of the first cultures of the Americas. As 'experts' many anthropologists have neither respected the endogenous knowledge systems and community contributions of Native Americans (or other indigenous peoples) nor addressed the intended and unintended impacts of anthropological research on those communities. Some anthropologists now acknowledge the harms that have been caused by researchers in the discipline, but it remains the case that anthropology must explicitly address the need to change its ways*."

The same should be said for the treatment of Legacy African Americans. African-descended peoples on the continent of Africa and throughout the African Diasporas have also been historically maligned and neglected by the scientific community. Even in contemporary genomic studies it is rare to hear an emic perspective of the African American genomics interpretation. By emic, we are referring to its anthropological use in denoting an approach to the study or description of a particular language or culture in terms of its internal elements and their functioning rather than in terms of any existing external framework. For our purposes in this manuscript, each of the authors is a member of the African American community with extensive research in community engagement, historical narratives, biological anthropology, and genomics. Our perspectives are indigenous, internal to the culture, and emic capturing the sensitives and diversities within our population.

The importance of African-descended populations in genomic studies and the development of a truly global genomic database cannot be underestimated. Given the evolutionary origins of humanity in Africa, we have long argued that the various state-sponsored human genome projects should have long ago focused on the genetics of recent African descendants to adequately reflect a more plausible template for our species (Kararach et al., 2011). A quarter of a century later, we still lack an adequate African-centered database for our species. Genomic research in Africa has a long way to go and genomic research among African Americans should be more advanced than it is presently, given the physical proximity and accessibility of this segment of American society. Researchers working in Africa have only studied between 5,000 and 10,000 whole genomes from the continent, compared with as many as 1 million whole genomes worldwide. Africa has received less than 1% of the global investment in genomics research and clinical studies. Genomics studies in Africa could contribute significantly to research worldwide in understanding our species since all our lineages ultimately trace back to Africa where *Homo sapiens* emerged some 300,000 years ago. Even those human lineages who left continental Africa over the past 80,000 years ago and spread across the planet carry only a subset of human genomic diversity. As a result of this evolutionary history, the people of Africa today carry more genetic diversity than those of any other continent. There are segments of human genome that can only be studied in Africans since these are the only populations within in which these unique sequences and genomic components are found.

Furthermore, populations of recent African descent are a growing segment of the world community, and these populations tend to be younger, so hopefully, African-descended individuals and communities will be around longer to benefit from todays and tomorrow's genomic innovations. In 1950 the population of Africa was 177 million and it grew 7.6 times to more than 1.341 billion in 2020. Africa is the continent with the youngest population worldwide. As of 2021, around 40 percent of the population is aged 15 years and younger, compared to a global average of 26 percent (Micheletti et al., 2020). Africa is quickly recovering from the destructive population losses associated with centuries of extractive enslavements facilitated by wars, exploitative colonialization by European and Arab powers, and years of local political mismanagement precipitated by low educational levels.

Legacy African Americans are not a genomic substitute for continental Africans as much autochthonic continental African genomic variation was lost among African Americans during the genetic bottlenecks of the transatlantic Middle Passage, the subsequent ravages of American Slavery, and the generations of forced gene flow with non-Africans. Instead, the justification for studying the genomics of Legacy African Americans stands independent and yet is connected to the need for comprehensive studies of African genetic diversity. In African Americans we have the unique opportunity as researchers to study the effects of well-specified gene-environment interactions on a historically socially restricted population that represents an amalgamation of West, West Central, and Southeast African peoples with modest gene flow from select non-African groups, primarily North Atlantic and Iberian Europeans and eastern Native American peoples.

The mobility of these early enslaved Africans was extremely circumscribed, largely following the forced migrations to North America. Countering this lack of geospatial movement was the fact that enslaved Africans represented, from the start, a broad array of geographically and culturally distinct African peoples. Initially these diverse Africans sorted themselves by their original African ethnic groups or their closest affiliates on the African continent. The initial retention of original identity provided a template for resistance among the survivors and their immediate descendants (e.g., the nearly constant slave rebellions and uprisings were often organized and implemented along African ethnic affiliation) and sexual selection (e.g., especially female-based mate selection may have

been based on African ethnicity and religious preference). The effectiveness of this self-sorting was density dependent. Initially we observe genomic aggregates based upon original ethnicity, but these aggregates were strongly discouraged by slave honors because of the enhanced potential for rebellion mentioned previously. Where there were larger numbers of enslaved Africans, such as the big plantations of the Southeastern and Mid-Atlantic regions of North America and in the urban areas of the American colonies such as New Amsterdam/New York City, African genomic integrity and cultural preferences could be retained longer and more cohesively. Over time, however, within the context of institutionalized, multi-generational enslavement, self-identities were transformed, and the original African ethnic affiliations gave way to new localized identities. This is the genesis of the many microethnic communities of African Americans that today can be found throughout the homeland and satellite territories of the African-descended peoples of the Americas. African Americans follow this same generalized population biology pattern of initial fusion followed by transformation and subsequent fission.

Of the limited number of comprehensive genomic studies done on African Americans, we can already see the promise of genomics to reveal major insights. A major DNA study (Loshin, 2002) recently shed new light on the fates of the more than 12.5 million Africans who were enslaved and traded to the Americas between 1,515 and the mid-19th century. More than 50,000 people took part in the study, which was able to identify more details of the "genetic impact" the trade has had on present-day populations in the Americas. The study laid bare the consequences of rape, maltreatment, disease, and racism. More than 2 million of the enslaved men, women, and children died enroute to the Americas. But the interpretation of the results in this major paper were ahistorical and overemphasized the presumed genetic affinities of African Americans to modern day Nigeria (Jackson, 2021).

Despite the errors, if African Americans genomic studies can be a rich source of insight into human evolutionary biology and evolutionary medicine, who should own the resulting data? The question of ownership of genomic data is fraught with cultural nuance and interpretation. Data ownership refers to both the possession of and responsibility for information as ownership implies power as well as control (Githaiga, 2021).

For African-descended populations, there is no single cultural mandate among the indigenous peoples of Africa. For example, on the question of land ownership, indeed, the East African Community (Kenya, Tanzania, Rwanda, Burundi, and Uganda) is currently struggling with contentious traditional cultural perceptions of land that have defined land ownership, use and access (Stokstad, 2019). Genomic variation is a valuable resource. So, perhaps it is more analogous to trees. The oldest dictums from a collective of East African ethnic groups suggests that whoever plants a tree, owns that tree and the products of that tree (e.g., the fruit, the oil, the sap, the lumber). Even with changes in land ownership, the tree belongs to whomever planted it and his or her descendants. This is an appropriate metaphor for the control of genomic data generated in the process of biomedical research and ancient DNA studies. The data clearly stay with the population of origin and their descendants; they own the products of their ancestral trees. Descendant communities must also be engaged in the analysis and interpretation of these data. While their ownership does not preclude non-indigenous access to data, the lines of responsibility must be grounded in the African American community. The past bio-colonial paradigm of external ownership of African American genomic resources should be rejected.

Within traditional Africa, communities are generally structured hierarchically such that their organizational structure serves somewhat as a buffer against genetic exploitation. And yet, African genomic studies here have too frequently been characterized by ethical dumping, in-and-out helicopter science, and over extrapolation of limited data by Western scientists with few ties to the local communities. Researchers gathered samples with scant regard for informed consent and without giving back information and other resources to the communities they studied. Outside of structured communities in Africa, the threat of genetic exploitation was expected to be protected against by local governments. These protections have clearly not been fully effective, however. A recent prominent example has been the United Kingdom.'s Wellcome Sanger Institute. Here, whistleblowers in 2020 privately accused Sanger of commercializing a gene chip without proper legal agreements with partner institutions and adequate informed consent of the hundreds of African people whose donated DNA was used to develop the chip (H3Africa, 2021). The institute confirmed that it did not commercialize the chips or profit from them but admitted that its relationship with some African partners has been "disrupted." Stellenbosch University in South Africa has demanded that Sanger return these samples. Sanger's mishandling of this extensive genomic sampling effort will likely contribute to the ongoing erosion of trust between researchers and diverse African people, setting back genomic research that could have been of benefit to Africans and their recent descendants. This controversy with a major genome research center will inevitably retard the study of African genomics because it will amplify the existing distrust between African communities and the Western scientific establishment. However, Africans have begun to initiate their own studies, aided and inspired substantially through the Human Heredity and Health in Africa (H3Africa) Initative (Jegede, 2009) led by Charles Nohuoma Rotimi who is the Director of the Trans-National Institutes of Health (NIH) center for research in genomics and global health. It is these initiatives among both Africans and African Americans that will provide the best protection against a continuation of past genomic abuses (Thompson et al., 2003; Pellegrino et al., 2007).

Additionally, the development and expansion of scientific expertise among Africans and African Americans in the genomic sciences will allow the development of significant capacity building within these segments of the scientific community and the development of trust with the larger social and cultural communities from which these new scientists have emerged. True informed consent can only come from a foundation of trust based on correct understanding. Trust is built on shared experiences, shared expectations, the anticipation of predictable outcomes, and is a central part of all human relationships. Informed consent emanates from an educated understanding of the issues at hand, an awareness of the limitations of the technology in use, an appreciation of the meaning of the results generated by the research, and past evidence of mutual goodwill among the

researchers and the researched. The specifics of informed consent will vary across the range of a species, indeed across the range of a stratified subset of the species. Among African Americans, informed consent may vary across North America since the perceptions of key cultural components also diverge regionally. For example, the recognition of the rights of the dead and the veneration of ancestors vary across the geospatial range of African Americans. In African American cultures with strong African retentions, such as the Gullah Geechee peoples, the veneration of ancestors is strong and while augmented by a belief in a supreme being, prayers and/or sacrifices are also offered to the ancestors who may be conceived as minor deities. In these communities, the disposition of skeletal and dental remains, tissue samples, and DNA samples may take on additional significance. Only through careful ethnographic inquiry and structured survey methods (e.g., the collection of qualitative and quantitative data from the actual African American communities of relevance) can we begin to document the nuance of diverse perspectives evident among African Americans with respect to genomic studies. In spite of the regional substructure among African Americans, there does exist a *"collective cast of mind."* (***Cited in*** (Wolinetz and Collins, 2020)) on the many issues that determine what is collectively valued, who the people consider themselves to be, what priorities define them as to who they are, and how them perceive themselves in the larger society. Without these data providing an authentic and collective voice of the people, researchers are not only sampling blindly and magnifying disparities, but they are denying African Americans the autonomy as laid out in western ethical principles (*see* (Sanders, 2021)).

The troubling victimization and exploitative history of Legacy African Americans by the early biomedical and genomic science studies of the United States lays a challenging foundation for ethical future studies. Researchers must be even more careful in acquiring and documenting fully informed consent from African American individuals and communities and providing any requested feedback on the research results and needed educational opportunities. As the African American community collectively becomes more astute as to the nature of scientific research, additional ethical requirements will emerge, particularly for genomic studies. For example, the technological innovation of CRISPR Cas9 (clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 now permits genome editing (also known as gene editing) giving scientists the ability to directly manipulate an organism's DNA. In 2014, one of the first cases of applying this technology to humans was the editing of the genome an African American with sickle cell anemia (Frangoul et al., 2021). This disease afflicts millions of people around the world, most of them of African descent. Some 100,000 African Americans are afflicted with the disease. After 6 years of work, that experimental treatment was approved for clinical trials by the United States Food and Drug Administration, enabling the first tests in humans of a CRISPR-based therapy to directly correct the mutation in the beta-globin gene responsible for sickle cell disease (Graves et al., 2022). Yet, the application of CRISPR cas9 also reduces population variation, which, according to evolutionary theory, increases a populations vulnerability to extinction. As CRISPR-based interventions become more widespread and of public health significance, the ethical and evolutionary implications of diminished population genomic variability in the quest for immediate improvements in individual health will have to be reconciled. Undoubtedly, African American communities will figure prominently in these discussions because of the historical legacy of western science seeking pathology (in the context of disease alleles) in Black bodies.

Clearly, the larger scientific community has an obligation to promote researchers from underrepresented communities at all levels of genomic sciences. This is, in fact, the best response to past wrongs, and the strongest deterrent against future ethical abuses. Recently, Graves and others (Graves and Goodman, 2021). called for a new agenda to address inequality in science. In this call, they stressed the need to attract individuals who have been historically excluded from participation in science and highlighted the importance of engaging in substantial work to overcome the longstanding obstacles to their full participation. This call cannot be overemphasized: multidimensional African American involvement in the genomic sciences is essential to make up for the current deficiencies in the global database and, just as importantly, to rectify the inadequacies in a comprehensive understanding of the genomic ramifications the African American experience in North America. Accurate, historical and culturally-contexed interpretations of the genomic data are as important as the raw genomic data themselves. In fact, to have the latter without the former provides little good for the African American population. In the authors experiences at Howard University, we have witnessed the value of interdisciplinary input in genomic science interpretation. We also have had the firsthand opportunity to work over a number of years with the two case studies presented below, the New York African Burial Ground and the Gullah Geechee peoples of the Carolina Lowcountry, evaluating both from emic perspectives.

Origins of African American mistrust in medicine and its consequences for genomic studies

North American patterns of institutionalized racism, state sponsored segregation, and social disenfranchisement in genomics are reflected in the historical medical practices of the country. Thus, the patterns of inequality remain a tenacious part of contemporary research practices and perceptions. Concepts such as race, ancestry, genetics, access, equity, equality, and medicine are intertwined and intractably interconnected due to the pervasive historical pattern of exploiting race as a biological construct (***see*** (Washington, 2006)). In *Medical Apartheid*, Washington (Thompson et al., 2003) describes the dehumanizing processing of enslaved Africans and their African American descendants upon their arrival in the Americas resulting from the transatlantic and domestic trades as they were sold to new "owners". Inadequate personal privacy, lack of sanitation, overcrowding, stark nutritional deprivations, and other detrimental public health conditions for enslaved Africans and their African American descendants meant enhanced exposures to infectious diseases from Europe, the Americas, and Africa, compounded by the disorders of nutritional deficiencies, the psychological and physical traumas of enslavement, and the

enslaved persons preemptive status as experimental models for early biomedical studies. This was done without the documented consent whatsoever of participants and these studies were enacted without the researchers understanding for or appreciation of the ancestral backgrounds or population substructure of African Americans. Black bodies were poked and prodded, surgeries were performed without available anesthesia, and known therapeutic medications were withheld. Simultaneously, non-traditional and herbal based medicinal practices were banned in Legacy African American communities. Stories of the "strength" and "lack of pain" experienced by African American women in childbirth plague their level of care in Labor and Delivery wards today. If enslaved African Americans complained about their aliments, these nascent physician-scientists responded according to the directions of the plantation owner whose goals were consistently to maximize their economic profits. This resulted in veterinarians "practicing" on humans and harmful "quick-fixes" done more often than necessary. Early experimental studies on exploited, enslaved, and newly freed African Americans were used to bolster tainted theories about European and European American supremacy in intellect and humanity and have set the historical template for the ethical challenges we currently face in studying the genetics of these continually marginalized communities. The prejudices and beliefs of this historical time has prevailing implications, even unconsciously in contemporary western medical spaces.

The mechanisms that have contributed to the marginalization of Legacy African Americans and their descendants, the importance of performing ethically responsible research on underrepresented populations, and the consequences of performing more inclusive, unbiased research on historic and contemporary African Americans emerge directly from the case studies we present.

Learning more about the genetics of historic populations, particularly, those buried in the NYABG helps us better understand the genetic identities of free and enslaved Africans, genetic adaptation due to the world's most extensive forced migration pressures, and genetic diseases that affected a historical population. In addition, increased knowledge of historic African American genomics allows researchers to comprehend better the genomics of living African Americans. Illuminating the genomics of African Americans is essential for several reasons such as: 1) it provides a multi-dimensional sense of identity, genomic and ancestral, that was severed by the Transatlantic Slave Trade, 2) it reveals the diversity within continental Africans, ultimately contributing to a greater understanding of all humankind and 3) it contributes to the paucity of African descended peoples in genomic databases to be used by medical professionals to make more informed diagnoses and treatment plans as we move into the age of precision medicine.

A premier concern in exploring these insights is ensuring ideal conditions (financial, ethical, and legal) are met to study African American genomic research appropriately. First, funding agencies must see the value in studying African and African American populations with an inclusive benefit for them and their descendants. Ethically responsible research to respectfully study underrepresented groups must become standard practice. Finally, legally, protections must be set to ensure the safeguarding of African American biological samples, remains, and genetic data (Jackson et al., 2021).

The need to protect African bodies was proven necessary upon the arrival of the first enslaved Africans to the United States based on the understanding that chattel slavery was dehumanizing and immoral. The need to protect African bodies from illegal biological research was a simultaneous necessity as many were purchased for the sole purpose of medical experimentation to advance the reputation and career of the purchasers (Thompson et al., 2003). An early example of using African bodies against their will and exploiting biological processes for financial and economic gain is in the work of J. Marion Sims, the "Father of Modern Gynecology" during the mid-18th century. He was praised in the medical world for his advancements in vesicovaginal fistula treatment and the first gallbladder surgery, which he developed and practiced on enslaved African women. However, it was not until recently that years of controversy stemming from Sims' ethical practices around discovering these advancements through his unorthodox experimentation on enslaved women led to a change. While there was no compensation for African Americans, retribution came in 2018 when New York City finally removed his statue from Central Park across from the New York Academy of Medicine (Walloo, 2018). Another example involves Georgia physician W. H. Robert and his inclination to amputate the limbs of enslaved Africans for minor injuries as demonstrations for medical students. He believed that students should "hesitate much less to remove a limb . . . , if he be slave, than if he be a free man, and especially a white man." This advice was based on Robert's observation that the surgical pain felt by an enslaved person was negligible, minor compared to what a white man facing the procedure would feel (Thompson et al., 2003). The idea that people of African descent do not possess the capacity to feel pain at the same intensity as white people still resonates throughout the medical industry today. Studies found that when a Black person enters an emergency room with pain like a broken bone and then a white person enters an emergency room with the same ailment, the Black person will receive a lesser dosage and even sometimes an inferior treatment. A 2016 survey of 222 white medical students and residents revealed racial bias in pain perception and accuracy of treatment, including less effective pain-relieving options (Hoffman et al., 2016b) for African Americans. Notions such as this are the basis for large-scale socio-economic crises, like the opioid epidemic.

Scientific research on African American remains

As identified in the studies of the New York African Burial Ground, just as enslaved Africans were controlled during their lives, European enslavers and public officials carried over this control even after their death. The need to protect African remains became necessary the moment they were buried. Misusing African remains has been demonstrated across medical colleges in the United States during the late 1700s and 1800s (Shultz, 2005; Royes, 2020). Employees of medical colleges, medical students,

and instructors would illegally dig up the cadavers of African Americans for anatomy instruction. The remains were used without the knowledge or permission of the person or their living relatives. The bones were never replaced after their teaching purpose was fulfilled. The affected families were never compensated. Furthermore, the bones of the unearthed individuals were never acknowledged for their contribution to scientific advancement (Thompson et al., 2003).

The Medical College of Georgia's (MCG) participation in "grave robbing" is of relevance. In 1989, a construction project renovating the Old Medical School building uncovered an estimated 9,000 human bones (350–450 people) buried in the basement. Most of the remains were taken from a predominantly African American cemetery, Cedar Grove, years before dissection of bodies became legal in 1887 (Taylor, 2019). However, even in this blatantly illegal and morally corrupt act of stealing bodies, the MCG did little more than recognize their predatory past. Only by revisiting the MCG discovery (along with other exploitative investigations of Black bodies and mishandling of their remains) and noting where more appropriate, respectful, and ethically responsible actions could have been taken can we truly understand the unfortunate foundations of the United States medical industry.

The challenge of studying historically underrepresented populations, particularly African Americans, is that their existence (in life and death) has been undervalued. As we have seen throughout history, if a group is undervalued, there is less investment for scientific researchers and physicians to benefit that group. Benefits include but are not limited to using informed consent (by researchers and medical professionals), allowing individuals to make autonomous decisions about their medical procedures, receiving medical care using the same methods that have been developed with the reluctant participation of enslaved Africans, and assured protection for burials from graverobbers or overzealous medical students. Unfortunately, the limited investment in African American research results in a failure to learn all we can about the genomic makeup of an underrepresented group in scientific and medical research. Further, because the limitation stunts our understanding of the genomic variation and diversity in African descended peoples, the population whose origin is located on the same continent as the inception of the*Homo*species, we fail to learn all that we can about the entire human population.

Absence of African genomic data in global databases

There are exceptions to the undervalued condition where historically marginalized groups, in this case, African Americans, are commoditized for their biological genomic data. Usually, these exceptions occur when research is performed to benefit European researchers and patients. An example of this exception is seen in the increasing thirst of commercial DNA testing companies to enrich their databases (Jaiswal and Halkitis, 2019). The origins of American medicine and the direction of medical practice are driving factors for inequities in our healthcare system and scientific research. As researchers work to expose, address, and dismantle how deeply entrenched biases have shaped scientific research and medicine, we are forced to consider how we presently deal with race, access, and health disparities. The reluctance of many African Americans to engage with the American medical system stems from a generational pattern of historical mistrust of the system and its founders (Gamble, 1993; Suite et al., 2007; Sirugo et al., 2019). We are approaching a fork in the road, where if researchers continue down the current path, where African descended people make up roughly 2% of global genomic database contributions (Popejoy and Fullerton, 2016), we will reach a point where African Americans are exponentially lagging (even more than the present status) in genomic research regarding health outcomes and the potential for personalized medicine applications. The large gap between the number of European participants in genomic databases and all other groups results from historical, cultural, scientific, and logistical factors sustaining bias in genomic research (Atutornu et al., 2022). Genome-wide association studies (GWAS) surveys show that over 70% of samples come from the United States, Iceland, and the United Kingdom. Choosing the path less followed means embarking upon a new Frontier where geographically and ethnically diverse genomic databases serve as an enriched reservoir for more accurate and less biased scientific research. Human genomic diversity between African genomes and the rest of the world results in differences between the variants associated with specific disorders and genes, making it more challenging to find the link between genetic variants and disease in African descended peoples. This challenge means that causal links between variants and disease cannot be trusted in medicine if the data upon which the diagnosis is formulated does not include populations from diverse ancestral backgrounds (Coles and Mensah, 2017). If an adjustment to this new path is not made, African Americans will continue to exponentially lag other groups in the race to precision medicine, or worse, be given the wrong genetic diagnosis or risk profile for disease. They will continue to be disadvantaged in genomic research opportunities leading to better overall health and access to personalized medicine applications, gene therapies, and pharmacogenomic benefits (Atutornu et al., 2022).

The historical mistrust between the African American community and the healthcare industry is a crucial factor contributing to missing data in genomic databases (Coles and Mensah, 2017). Tackling this predicament requires the continued rebuilding of confidence at every level of healthcare to demonstrate its investment in the lives of African Americans. While an exact solution is unclear, we hypothesize that once developed, it will take years of application to rebuild trust among African Americans. Researchers are working to combat the paucity of diverse data among living African Americans in genomic databases through initiatives such as the H3Africa consortium (Bentley et al., 2020) and the *All of Us* research program (Department of Health and Human Services, 2019). Others are working on grasping a more robust understanding of African American genetics through studying African American remains. One way to combat the missing data issue is by analyzing historic African American genomes. With the permission of the descendants of these buried populations, researchers can address the dire need to enrich genomic databases in two ways. The first is to increase the numbers of African descended genomes in the databases, and the second is by widening the breadth of information that can be learned about a population by studying individuals who lived hundreds of years ago. The relatedness of individuals in a population coalesces as

you travel backward in time and thus gives researchers a broader scope of the genetics of living descendants without needing their samples directly. Genomic data from historic remains gives us insight into the health disparities, genetic variation, and disease susceptibilities of living Legacy African Americans. Research on historical remains provides a window into the genetics of living African Americans circumventing this historical mistrust and fear to ensure a future for access to precision medicine for this underrepresented group. Pushing human remains research forward, we at Howard University set out to observe human evidence in NYABG burial soil samples that have been buried for four hundred-years.

Ethical influences on genomic testing of African Americans

Prior to the inclusion of ethical principles in the routine training of physicians and scientists, enslaved and newly freed African Americans were disproportionately represented in unregulated experimental studies and were the targets of eugenic hypotheses. Once application of the ethical principles of autonomy, informed consent, privacy/confidentiality, beneficence, nonmaleficence, and justice became commonplace in western science, the collection of global genomic databases became overwhelmingly comprised of the genomic data of peoples of North Atlantic European ancestry. This current fact presents continuing limitations for all other (non-European) peoples and the extent of their deficit is proportional to their degree of difference from this North Atlantic European standard. The impact of the underrepresentation is particularly acute for associated health implications, for inadequate genomic and medical research lays a foundation for the perpetuation and amplification of current health disparities among the most disenfranchised. Populations of recent African descent, for example, have greater genetic variation when compared with other non-African populations. African Americans are an accessible population for capturing a proportion of the genomic diversity of Africa. The failure to include African Americans in genomic studies may lead to increased health disparities (Jackson, 1997); what is not known cannot be properly addressed, and vital ancestral history will continue to be missed in these communities. While race is not a genetically meaningful category, its social ramifications continue to impact biology through the enactment of racist policies and practices which result in inequities in areas such as healthcare. As we learn more about the fine mapping and interactions of ancestral origins and their correlated disease risks, researchers will be restricted in their capacity to address health disparities, evaluate appropriate applications for precision medicine, and understand the broad landscape of the human genome with such a limited and skewed global genomic database. While these limitations were recognized 25 years ago (Jackson, 1997; Jackson, 1998), the genomic community has been slow to address this equity issue.

Conclusion

Given these core issues, how do we forge a research agenda that addresses the expanding marginality of underrepresented groups such as Legacy African Americans (and African-descended peoples in general) (Rogers and Lange, 2013) in the face of rapid technological advances in genomics and the increasingly direct applications to genomics to clinical diagnostics and therapeutic intervention (e.g., CRISPR Cas9 gene therapy)? We posit that there is important urgency to address the current paucity of Legacy African American genomics specifically and African genomics in general. It is necessary to expand the scope and volume of inclusion for non-European populations to ensure equity in healthcare. Today, all humans alive on Earth share a common ancestor who can be traced back to continental Africa (Cann et al., 1987). Human residence has been the longest in Africa and the original population sizes were larger than elsewhere. Additionally, Africa alone comprises at least 11 ancestral groups compared to 12 ancestral groups in the rest of the world (Kwok, 2009). With the deepest evolutionary history and the greatest diversity, African genomes can tell us more about the health and existence of humankind than any other population. Genomic databases must be enriched with African descended genomes to paint the most accurate picture of who we are as a species. Perhaps efforts should be made to refine the content of current genomic databases to represent the entire human population accurately. As a step toward parity in genomics, what if databases were 90% African and 10% all other populations? This formulation would be a means to recalibrate our assessments to make them evolutionarily more profound and reflective of a broader cross-section of our species diversity. Such enhanced representativeness is also needed for the future endeavors of our species, particularly genomic modifications that will be needed to make human life on other planets sustainable. We may already have among our species the allelic variants and epigenetic markers that could augment our future extraterrestrial existence.

The most insidious shortcoming of missing genomic data from non-European populations however is the harm it poses to the health and survival of non-European peoples. Due to the paucity of genomic data on African American populations, reportedly "rare variants" do not accurately reflect the overall data but are a product of the bias due to a lack of diversity in genomic research. The absence of data leads to misdiagnosis of the origin of the disease or disorder. Additionally, if non-European populations are not adequately represented in genomic research, they cannot access its benefits, such as gene therapy and precision medicine, including pharmacogenomics that they contribute to as taxpayers. All in all, adequate representation in genomic databases translates to better and more equitable health outcomes and preventative treatment for all people. The rationale for inclusion is clear and the mechanisms needed to ensure that this inclusion is ethical are feasible. What we now lack is the will to implement these important innovations.

Data availability statement

22 publicly available databases were evaluated to create the metadatabase used in this study. Links to these individual databases are provided in the text. Further inquiries about the metadatabase can be directed to the corresponding author.

Ethics statement

The studies presented involved previously published data that were approved by Howard University's Institutional Review Board. All participants provided written informed consent at the time data were collected.

Author contributions

All authors contributed to the article and approved the submitted version. The concept and organization of the review paper was by FJ, the New York African Burial Ground case study by CC, and the Gullah Geechee case study by JC.

Acknowledgments

We thank the *All of Us* Researcher Academy (NIH) and RTI for support during the revisions of this review. We also appreciate the quantitative insights of QuadGrid Data Lab in the development of the project's metadatabase.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

AABA Code of Ethics. Available at; https://bioanth.org/documents/3/ethics.pdf#:~:text=The%20American%20Association%20of%20Physical%20Anthropologists%20%28AAPA%29%20does,ethical%20framework%2C%20as%20they%20engage%20in%20their%20work (Accessed May 29, 2023).

Abel, S., and Schroeder, H. (2020). From country marks to DNA markers: The genomic turn in the reconstruction of african identities. *Curr. Anthropol.* 61 (22), S198–S209. doi:10.1086/709550

Accessible Archives Accessible archives database. Available at: https://www.accessible.com/accessible/preLog.

African American Home Movie Archive. Available at: https://www.aahma.org/ (Accessed May 29, 2023).

Afro-American Genealogical Research (2020). Introduction 2020. Available at: https://guides.loc.gov/african-american-genealogical-research (Accessed September 14, 2020).

American Freedmen's Inquiry Commission. Available at; https://researchworks.oclc.org/archivegrid/data/612775411 (Accessed May 29, 2023).

AnthroSource Database. Available at: https://anthrosource.onlinelibrary.wiley.com/ (Accessed May 29, 2023).

Atutornu, J., Milne, R., Costa, A., Patch, C., and Middleton, A. (2022). Towards equitable and trustworthy genomics research. *EBioMedicine* 76, 103879. doi:10.1016/j.ebiom.2022.103879

Baharian, S., Barakatt, M., Gignoux, C. R., Shringarpure, S., Errington, J., Blot, W. J., et al. (2016). The Great migration and african-American genomic diversity. *PLoS Genet.* 12 (5), e1006059. doi:10.1371/journal.pgen.1006059

Bailey, C. (1995). *Personal communication*. Sapelo Island, GA: University of Georgia.

Barker, L. E., Kirtland, K. A., Gregg, E. W., Geiss, L. S., and Thompson, T. J. (2011). Geographic distribution of diagnosed diabetes in the U.S.: A diabetes belt. *Am. J. Prev. Med.* 40 (4), 434–439. doi:10.1016/j.amepre.2010.12.019

Bentley, A. R., Callier, S. L., and Rotimi, C. N. (2020). Evaluating the promise of inclusion of African ancestry populations in genomics. *npj Genomic Med.* 5, 1–9. doi:10.1038/s41525-019-0111-x

Berlin, I. (2010). *The making of african America: The four Great migrations*. New York, NY: Viking Press.

Caldwell, J., and Jackson, F. L. C. (2021). Evolutionary perspectives on african North American genetic diversity: Origins and prospects for future investigations. *Evol. Anthropol.* 30, 242–252. doi:10.1002/evan.21910

Cann, R., Stoneking, M., and Wilson, A. (1987). Mitochondrial DNA and human evolution. *Nature* 325, 31–36. doi:10.1038/325031a0

Clinton, C. (2021). "Elemental and bacterial analysis of 17th and 18th Century New York African Burial Ground grave soil samples with bioethical implications for scientific research," *PhD dissertation* (Washington: Howard University).

Clinton, C. K., and Jackson, F. L. C. (2021). Historical overview, current research, and emerging bioethical guidelines in researching the New York African burial ground. *Am. J. Phys. Anthropol.* 175, 339–349. doi:10.1002/ajpa.24171

Coles, E., and Mensah, G. A. (2017). Geography of genetics and genomics research funding in Africa. *Glob. heart* 12 (2), 173–176. doi:10.1016/j.gheart.2017.07.001

Department of Health and Human Services (2019). All of us *research program*.

Eltis, D., Engerman, S. L., Drescher, S., and Richardson, D. (2017). *The cambridge world history of slavery*. Cambridge UK: Cambridge University Press.

Eltis, D., and Richardson, D. (2015). *Atlas of the transatlantic slave trade*. New Haven: Yale University Press.

Ely, B., Wilson, J. L., Jackson, F., and Jackson, B. A. (2006). African American mitochondrial DNAs often match mtDNAs found in multiple African ethnic groups. *BMC Biol.* 4, 34. doi:10.1186/1741-7007-4-34

Emory university Emory University. Available at: https://proxy.library.emory.edu/login?url=http://search.ebscohost.com/login.aspx?authtype=ip,uid&profile=ehost&defaultdb=sih.

Esenwa, C., Tshiswaka, D. I., Gebregziabher, M., and Ovbiagele, B. (2018). Historical slavery and modern-day stroke mortality in the United States stroke belt. *Stroke* 49 (2), 465–469. doi:10.1161/STROKEAHA.117.020169

Fatumo, S., Chikowore, T., Choudhury, A., Ayub, M., Martin, A. R., and Kuchenbaecker, K. (2022). A roadmap to increase diversity in genomic studies. *Nat. Med.* 28 (2), 243–250. doi:10.1038/s41591-021-01672-4

Fleskes, R. E., Ofunniyin, A. A., Gilmore, J. K., Poplin, E., Abel, S. M., Bueschgen, W. D., et al. (2021). Ancestry, health, and lived experiences of enslaved Africans in 18th century Charleston: An osteobiographical analysis. *Am. J. Phys. Anthropol.* 175 (1), 3–24. doi:10.1002/ajpa.24149

Fortes-Lima, C., and Verdu, P. (2021). Anthropological genetics perspectives on the transatlantic slave trade. *Hum. Mol. Genet.* 30, R79–R87. doi:10.1093/hmg/ddaa271

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y-S., Domm, J., Eustace, B. K., et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. *N. Engl. J. Med.* 384, 252–260. doi:10.1056/NEJMoa2031054

Gamble, V. N. (1993). A legacy of distrust: African Americans and medical research. *Am. J. Prev. Med.* 9 (6), 35–38. doi:10.1016/s0749-3797(18)30664-0

GenBank. National Center for Biotechnology Information. Available at: https://www.ncbi.nlm.nih.gov/genbank/ (Accessed May 29, 2023).

1000 Genomes Project. The International Genome Sample Resource. Available at: https://www.internationalgenome.org/ (Accessed May 29, 2023).

Githaiga, N. (2021). Much ado about land in the EAC (online) institute for security studies. Available at: https://issafrica.org/iss-today/much-ado-about-land-in-the-eac (Accessed November 26, 2021).

Google Scholar Search Engine. Available at: https://scholar.google.com/ (Accessed May 29, 2023).

Graves, J. L., and Goodman, A. H. (2021). *Racism not race*. New York, NY: Columbia University Press, 320.

Graves, J. L., Jr, Kearney, M., Barabino, G., and Malcom, S. (2022). Inequality in science and the case for a new agenda. *Proc. Natl. Acad. Sci. U. S. A.* 119 (10), e2117831119. doi:10.1073/pnas.2117831119

Grice, E. A., and Segre, J. A. (2012). The human microbiome: Our second genome. *Annu. Rev. Genomics Hum. Genet.* 13, 151–170. doi:10.1146/annurev-genom-090711-163814

Grundy, S. (1998). Multifactorial causation of obesity: Implications for prevention. *Am. J. Clin. Nutr.* 67, 563S–72S. doi:10.1093/ajcn/67.3.563S

GSA. US general Services administration. Available at; https://www.gsa.gov/ (Accessed May 29, 2023).

Gultig, K. (2023). Puzzling our roots: What does diversity have to do with genetics and disease? Available at: https://www.vusci.org/puzzling-our-roots-what-does-diversity-have-to-do-with-genetics-and-disease/.

Gupta, A. (2021). Available at: https://www.americananthro.org/StayInformed/NewsDetail.aspx?ItemNumber=28239.

H3Africa (2021). H3Africa funds scientists, advances genomics, and fosters collaboration in Africa. Available at: https://h3africa.org/index.php/2021/12/06/h3africa-funds-scientists-advances-genomics-and-fosters-collaboration-in-africa/5 (Accessed December 3, 2021).

Hesburgh libraries Social sciences full Text (wilson Web). Available at: https://www.library.nd.edu/database/4qBOLHGnSwwM0ekI4SUkKC.

Hine, D. C., Hine, W. C., and Harrold, S. C. (2010). *African-American Odyssey*. 5th. New York, NY: Pearson Publications.

Hoffman, K. M., Trawalter, S., Axt, J. R., and Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proc. Natl. Acad. Sci. U. S. A.* 113 (16), 4296–4301. doi:10.1073/pnas.1516047113

Hoffman, K. M., Trawalter, S., Axt, J. R., and Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proc. Natl. Acad. Sci.* 113 (16), 4296–4301. doi:10.1073/pnas.1516047113

Holm, J. (1983). On the relationship of Gullah and Bahamian. *Am. Speech* 58 (4), 303–318. doi:10.2307/455145

Jackson, F., Clinton, C., and Graves, J. (2021). Genomics of african American remains — Limits must not compound inequity. *Nature* 596, 35. doi:10.1038/d41586-021-02129-x

Jackson, F., Jackson, L., and Jackson, Z. E. (2018). Developmental stage epigenetic modifications and clinical symptoms associated with the trauma and stress of enslavement and institutionalized racism. *J. Clin. Epigenetics* 4, 11. doi:10.21767/2472-1158.100096

Jackson, F. (2004). Human genetic variation and health: New assessment approaches based on ethnogenetic layering. *Br. Med. Bull.* 69, 215–235. doi:10.1093/bmb/ldh012

Jackson, F. L. C. (1997). Concerns and priorities in genetic studies: Insights from recent African American biohistory. *Set. Hall. Law Rev.* 27 (3), 951–970.

Jackson, F. L. C. (2008). Ethnogenetic Layering (EL): An alternative to the traditional race model in human variation and health disparity studies. *Ann. Hum. Biol.* 35 (2), 121–144. doi:10.1080/03014460801941752

Jackson, F. L. C. (1998). Scientific limitations and ethical ramifications of a non-representative Human Genome Project: African American response. *Sci. Eng. Ethics* 4, 155–170. doi:10.1007/s11948-998-0046-6

Jackson, F. L. C. (2021). So many Nigerians: Why is Nigeria overrepresented as the ancestral genetic homeland of legacy african North Americans? *Am. J. Hum. Genet.* 108 (1), 202–208. doi:10.1016/j.ajhg.2020.10.010

Jackson, L., Jackson, Z., and Jackson, F. (2018). Intergenerational resilience in response to the stress and trauma of enslavement and chronic exposure to institutionalized racism. *J. Clin. Epigenetics* 4, 15. doi:10.21767/2472-1158.1000100

Jaiswal, J., and Halkitis, P. N. (2019). Towards a more inclusive and dynamic understanding of medical mistrust informed by science. *Behav. Med.* 45 (2), 79–85. doi:10.1080/08964289.2019.1619511

Jegede, A. S. (2009). Culture and genetic screening in Africa. *Dev. World Bioeth.* doi:10.1111/j.1471-8847.2009.00259.x

Kararach, G., Hanson, K. T., and Leutier, F. A. (2011). Regional integration policies to support job creation for Africa's burgeoning youth population. *World J. Entrepreneursh. Manag. Sustain. Dev.* 7, 177–215. doi:10.1108/20425961201000035

Karp, D. N., Wolff, C. S., Wiebe, D. J., Branas, C. C., Carr, B. G., and Mullen, M. T. (2016). Reassessing the stroke belt: Using small area spatial statistics to identify clusters of high stroke mortality in the United States. *Stroke* 47 (7), 1939–1942. doi:10.1161/STROKEAHA.116.012997

Kramer, M. R., Black, N. C., Matthews, S. A., and James, S. A. (2017). The legacy of slavery and contemporary declines in heart disease mortality in the U.S. South. *SSM - Popul. Health* 3, 609–617. doi:10.1016/j.ssmph.2017.07.004

Kwok, R. (2009). Cellulosic ethanol hits roadblocks. *Nature* 461, 582–583. doi:10.1038/461582a

Laster, M., Shen, J. I., and Norris, K. C. (2018). Kidney disease among african Americans: A population perspective. *Am. J. Kidney Dis.* 72 (5), S3–S7. doi:10.1053/j.ajkd.2018.06.021

Loshin, D. (2002). Knowledge integrity: Data ownership (online website). Available at: http://www.datawarehouse.com/article/?articleid=3052 (Accessed November 26, 2021).

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., Daly., M. J., et al. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* 51, 584–591. doi:10.1038/s41588-019-0379-x

Matory, J. L. (2008). The illusion of isolation: The Gullah geechees and the political Economy of african culture in the Americas. *Comp. Stud. Soc. Cult.* 50 (4), 949–980. doi:10.1017/s0010417508000406

McAdams-DeMarco, M. A., Daubresse, M., Bae, S., Gross, A. L., Carlson, M. C., and Segev, D. L. (2018). Dementia, alzheimer's disease, and mortality after hemodialysis initiation. *Clin. J. Am. Soc. Nephrol.* 13 (9), 1339–1347. doi:10.2215/CJN.10150917

Micheletti, S. J., Bryc, K., Esselmann, S. G. A., Freyman, W. A., Moreno, M. E., Poznik, G. D., et al. (2020). Genetic consequences of the transatlantic slave trade in the Americas. *Am. J. Hum. Genet.* 107 (2), 265–277. doi:10.1016/j.ajhg.2020.06.012

National Archives Records of the continental and confederation Congresses and the constitutional convention. Available at; https://www.archives.gov/.

National Archives RG 233: Records of the United States House of Representatives. Available at; https://www.archives.gov/findingaid/stat/discovery/233 (Accessed May 29, 2023).

NCIB Reference Sequence NCIB reference sequence (RefSeq). Available at: https://www/mcbo/m;,/moj/gpvrefseq.

Oklahoma Historical Society. "All-Black towns," the encyclopedia of Oklahoma history and culture. Available at: https://www.okhistory.org/ (Accessed May 29, 2023).

OMIM Online mendelian inheritance in man (OMIM). Available at: https:www.omim.org/.

Opala, J. A. (2009). *The Gullah: Rice, slavery, and the Sierra Leone-American connection*. Charlottesville, Virginia: USIS Originally from the University of Virginia.

Open Data for Africa Open data for Africa. Available at: https://dataportal.opendataforafrica.org/.

Paris, P. J. (1995). *The spirituality of african peoples: The search for a common moral discourse*. Minneapolis: Fortress Press.

Pellegrino, E. D. (2007). "Introduction: Culture and bioethics: Where ethics and mores meet," in *African American bioethics: Culture, race, and identity*. Editors E. D. Pellegrino and L. J. Prograis (Washington: Georgetown University Press).

Popejoy, A. B., and Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature* 538 (7624), 161–164. doi:10.1038/538161a

Sociological Abstracts Database. Applied social sciences Index and Abstracts (ASSIA). Available at: https://about.proquest.com/en/products-services/socioabs-set-c/ (Accessed May 29, 2023).

Rogers, W., and Lange, M. M. (2013). Rethinking the vulnerability of minority populations in research. *Am. J. public health* 103 (12), 2141–2146. doi:10.2105/AJPH.2012.301200

Royes, D. (2020). "Years of medical abuse make Black Americans less likely to trust the coronavirus vaccine: Reckoning with our past is crucial to getting buy-in for the vaccine. The Washington Post," in *Made by history*.

Sanders, $ (2021). FDA approves first test of CRISPR to correct genetic defect causing sickle cell disease. Available at: https://news.berkeley.edu/2021/03/30/fda-approves-first-test-of-crispr-to-correct-genetic-defect-causing-sickle-cell-disease/ (Accessed March 30, 2021).

Scharff, D. P., Mathews, K. J., Jackson, P., Hoffsuemmer, J., Martin, E., and Edwards, D. (2010). More than tuskegee: Understanding mistrust about research participation. *J. health care poor underserved* 21 (3), 879–897. doi:10.1353/hpu.0.0323

Schneider, S. P., and Gonzalez, F. J. (2021). Racial resentment predicts eugenics support more robustly than genetic attributions. *Personality Individ. Differ.* 180, 110989. doi:10.1016/j.paid.2021.110989

Shultz, S. M. (2005). *Body snatching: The robbing of graves for the education of physicians in early nineteenth century America*. McFarland.

Sirugo, G., Williams, S. M., and Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell.* 177 (1), 26–31. doi:10.1016/j.cell.2019.02.048

Slave Voyages Database. TransAtlantic slave Voyages database. Available at: https://www.slavevoyages.org/voyage/database (Accessed May 29, 2023).

Sociological Abstracts Sociological Abstracts. Available at; https://proquest.libguides.com/socabs.

Stokstad, E. (2019). Genetics lab accused of misusing African DNA. *Science* 366, 555–556. doi:10.1126/science.366.6465.555

Suite, D. H., La Bril, R., Primm, A., and Harrison-Ross, P. (2007). Beyond misdiagnosis, misunderstanding and mistrust: Relevance of the historical perspective

in the medical and mental health treatment of people of color. *J. Natl. Med. Assoc.* 99 (8), 879–885.

Tang, M. X., Cross, P., Andrews, H., Jacobs, D. M., Small, S., Bell, K., et al. (2001). Incidence of AD in african-Americans, caribbean hispanics, and caucasians in northern manhattan. *Neurology* 56 (1), 49–56. doi:10.1212/wnl.56.1.49

Taylor, M. E. (2019). How U.S. colleges stole freshly-buried black corpses in the 1830s for study. Face2Face Africa. Available at: https://face2faceafrica.com/article/how-u-s-colleges-stole-freshly-buried-black-corpses-in-the 1830s-for-study (Accessed November 23, 2021).

The African American Biographical Database. Available at: http://aabd.chadwyck.com/ (Accessed May 29, 2023).

The Library of Congress Available at: https://www.loc.gov/ (Accessed May 29, 2023).

The World Bank Database (WHO). Available at: https://data.worldbank.org/ (Accessed May 29, 2023).

Thompson, H. S., Valdimarsdottir, H. B., Jandorf, L., and Redd, W. (2003). Perceived disadvantages and concerns about abuses of genetic testing for cancer risk: Differences across african American, latina and caucasian women. *Patient Educ. Couns.* 51 (3), 217–227. doi:10.1016/S0738-3991(02)00219-7

Walloo, K. (2018). Historical aspects of race and medicine: The case of J. Marion Sims. *JAMA.* 320 (15), 1529–1530. doi:10.1001/jama.2018.11944

Washington, H. A. (2006). *Medical Apartheid: The dark history of medical experimentation on Black Americans from colonial times to the present*. New York, NY: Doubleday Books.

Wolinetz, C. D., and Collins, F. S. (2020). Recognition of research participants' need for autonomy: Remembering the legacy of henrietta lacks. *JAMA* 324 (11), 1027–1028. doi:10.1001/jama.2020.15936

Zimmerman, K. D., Schurr, T. G., Chen, W. M., Nayak, U., Mychaleckyj, J. C., Quet, Q., et al. (2021). Genetic landscape of Gullah african Americans. *Am. J. Phys. Anthropol.* 175 (4), 905–919. doi:10.1002/ajpa.24333

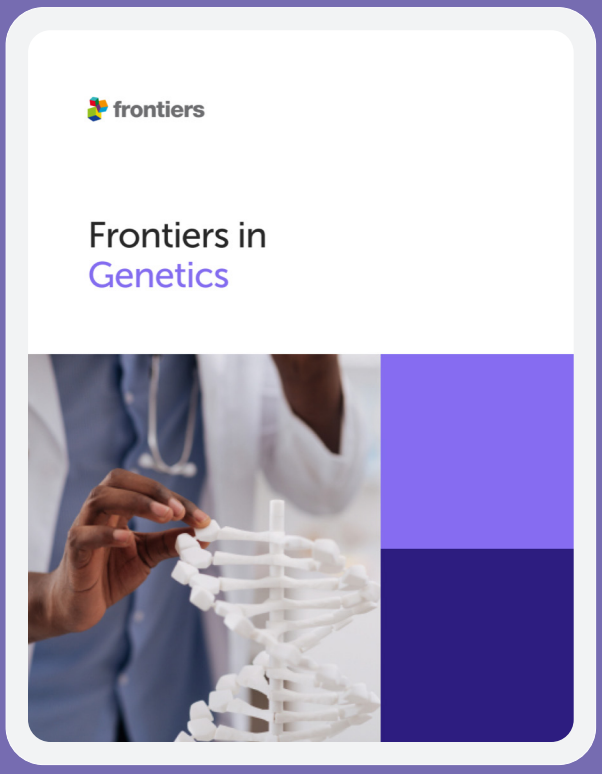
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact