

SCIENCE AND SOCIETY

Taking race out of human genetics

Engaging a century-long debate about the role of race in science

By Michael Yudell,^{1*} Dorothy Roberts,²
Rob DeSalle,³ Sarah Tishkoff²

In the wake of the sequencing of the human genome in the early 2000s, genome pioneers and social scientists alike called for an end to the use of race as a variable in genetic research (1, 2). Unfortunately, by some measures, the use of race as a biological category has increased in the postgenomic age (3). Although inconsistent definition and use has been a chief problem with the race concept, it has historically been used as a taxonomic categorization based on common hereditary traits (such as skin color) to elucidate the relationship between our ancestry and our genes. We believe the use of biological concepts of race in human genetic research—so disputed and so mired in confusion—is problematic at best and harmful at worst. It is time for biologists to find a better way.

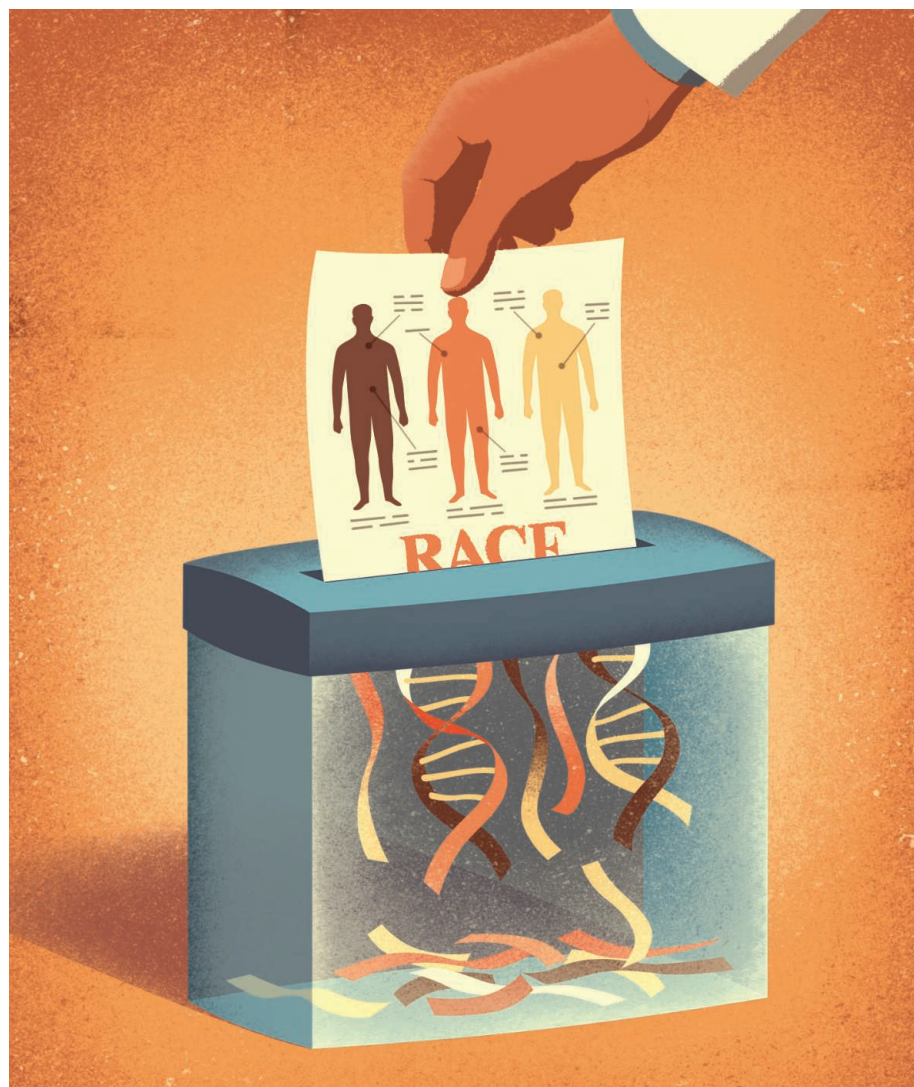
Racial research has a long and controversial history. At the turn of the 20th century, sociologist and civil rights leader W. E. B. Du Bois was the first to synthesize natural and social scientific research to conclude that the concept of race was not a scientific category. Contrary to the then-dominant view, Du Bois maintained that health disparities between blacks and whites stemmed from social, not biological, inequality (4). Evolutionary geneticist Theodosius Dobzhansky, whose work helped reimagine the race concept in the 1930s at the outset of the evolutionary synthesis, wrestled with many of the same problems modern biologists face when studying human populations—for example, how to define and sample populations and genes (5). For much of his career, Dobzhansky brushed aside criticism of the race concept, arguing that the problem with race was not its scientific use, but its nonscientific misuse. Over time, he grew disillusioned, concerned that scientific study of human diversity had “floundered in confusion and misunderstanding” (6). His transformation from defender to detractor of the race concept in biology still resonates.

Today, scientists continue to draw wildly different conclusions on the utility of the race

concept in biological research. Some have argued that relevant genetic information can be seen at the racial level (7) and that race is the best proxy we have for examining human genetic diversity (8, 9). Others have concluded that race is neither a relevant nor accurate way to understand or map human genetic diversity (10, 11). Still others have argued that race-based predictions in clinical settings, because of the heterogeneous nature of racial groups, are of questionable use (12), particularly as the prevalence of admixture increases across populations.

Several meetings and journal articles have called attention to a host of issues, which include (i) a proposed shift to “focus

on racism (i.e., social relations) rather than race (i.e., supposed innate biologic predisposition) in the interpretation of racial/ethnic ‘effects’” (13); (ii) a failure of scientists to distinguish between self-identified racial categories and assigned or assumed racial categories (14); and (iii) concern over “the haphazard use and reporting of racial/ethnic variables in genetic research” (15) and a need to justify use of racial categories relative to the research questions asked and methods used (6). Several academic journals have taken up this last concern and, with mixed success, have issued guidelines for use of race in research they publish (16). Despite these concerns, there have been no



¹Drexel University, Philadelphia, PA 19104, USA. ²University of Pennsylvania, Philadelphia, PA 19104, USA. ³American Museum of Natural History, New York, NY 10024, USA. *Corresponding author. E-mail: myudell@drexel.edu

systematic attempts to address these issues and the situation has worsened with the rise of large-scale genetic surveys that use race as a tool to stratify these data (17).

It is important to distinguish ancestry from a taxonomic notion such as race. Ancestry is a process-based concept, a statement about an individual's relationship to other individuals in their genealogical history; thus, it is a very personal understanding of one's genomic heritage. Race, on the other hand, is a pattern-based concept that has led scientists and laypersons alike to draw conclusions about hierarchical organization of humans, which connect an individual to a larger pre-conceived geographically circumscribed or socially constructed group.

Unlike earlier disagreements concerning race and biology, today's discussions generally lack clear ideological and political antipodes of "racist" and "nonracist." Most contemporary discussions about race among scientists concern examination of population-level differences between groups, with

"...the use of biological concepts of race in human genetic research...is problematic at best and harmful at worst."

the goal of understanding human evolutionary history, characterizing the frequency of traits within and between populations, and using an individual's self-identified ancestry to identify genetic risk factors of disease and to help determine the best course of medical treatments (6).

If this is what race in contemporary scientific and medical practice is about, then why should we be concerned? One reason is that phylogenetic and population genetic methods do not support a priori classifications of race, as expected for an interbreeding species like *Homo sapiens* (11, 18). As a result, racial assumptions are not the biological guideposts some believe them to be, as commonly defined racial groups are genetically heterogeneous and lack clear-cut genetic boundaries (10, 11). For example, hemoglobinopathies can be misdiagnosed because of the identification of sickle-cell as a "Black" disease and thalassemia as a "Mediterranean" disease (10). Cystic fibrosis is underdiagnosed in populations of African ancestry, because it is thought of as a "White" disease (19). Popular misinterpretations of the use of race in genetics also continue to fuel racist beliefs, so much so that, in 2014, a group of leading human population geneticists publicly refuted

claims about the genetic basis of social differences between races (20). Finally, the use of the race concept in genetics, an issue that has vexed natural and social scientists for more than a century, will not be obviated by new technologies. Although the low cost of next-generation sequencing has facilitated efforts to sequence hundreds of thousands of individuals, adding whole-genome sequences does not negate the fact that racial classifications do not make sense in terms of genetics.

More than five decades after Dobzhansky called on biologists to develop better methods for investigating human genetic diversity (21), biology remains stuck in a paradox that reflects Dobzhansky's own struggle with the race concept: both believing race to be a tool to elucidate human genetic diversity and believing that race is a poorly defined marker of that diversity and an imprecise proxy for the relation between ancestry and genetics. In an attempt to resolve this paradox and to improve study of human genetic diversity, we propose the following.

Scientific journals and professional societies should encourage use of terms like "ancestry" or "population" to describe human groupings in genetic studies and should require authors to clearly define how they are using such variables. It is preferable to refer to geographic ancestry, culture, socioeconomic status, and language, among other variables, depending on the questions being addressed, to untangle the complicated relationship between humans, their evolutionary history, and their health. Some have shown that substituting such terms for race changes nothing if the underlying racial thinking stays the same (22, 23). But language matters, and the scientific language of race has a considerable influence on how the public (which includes scientists) understands human diversity (24). We are not the first to call for change on this subject. But, to date, calls to rationalize the use of concepts in the study of human genetic diversity, particularly race, have been implemented only in a piecemeal and inconsistent fashion, which perpetuates ambiguity of the concept and makes sustained change unfeasible (16). Having journals rationalize the use of classificatory terminology in studying human genetic diversity would force scientists to clarify their use and would allow researchers to understand and interpret data across studies. It would help avoid confusing, inconsistent, and contradictory usage of such terms.

Phasing out racial terminology in biological sciences would send an important message to scientists and the public alike: Historical racial categories that are treated as natural and infused with notions of superiority and inferiority have no place in biology. We acknowledge that using race as a political or social category to study racism and

its biological effects, although fraught with challenges, remains necessary. Such research is important to understand how structural inequities and discrimination produce health disparities in socioculturally defined groups.

The U.S. National Academies of Sciences, Engineering, and Medicine should convene a panel of experts from biological sciences, social sciences, and humanities to recommend ways for research into human biological diversity to move past the use of race as a tool for classification in both laboratory and clinical research. Such an effort would bring stakeholders together for a simple goal: to improve the scientific study of human difference and commonality. The committee would be charged with examining current and historical usage of the race concept and ways current and future technology may improve the study of human genetic diversity; thus, they could take up Dobzhansky's challenge that "the problem that now faces the science of man [sic] is how to devise better methods for further observations that will give more meaningful results" (21). Regardless of where one stands on this issue, this is an opportunity to strengthen research by thinking more carefully about human genetic diversity. ■

REFERENCES AND NOTES

1. F. Collins, *Nat. Genet.* **36** (suppl.), S13 (2004).
2. M. W. Foster, R. R. Sharp, *Nat. Rev. Genet.* **5**, 790 (2004).
3. P. A. Chow-White, S. E. Green Jr., *Int. J. Commun.* **7**, 556 (2013).
4. W. E. B. Du Bois, *The Health and Physique of the Negro American* (Publ. no. 11, Atlanta Univ. Publications, Atlanta, GA, 1906).
5. T. Dobzhansky, *Genetics and the Origin of Species* (Columbia Univ. Press, New York, 1937).
6. M. Yudell, *Race Unmasked: Biology and Race in the 20th Century* (Columbia Univ. Press, New York, 2014).
7. E. G. Burchard et al., *N. Engl. J. Med.* **348**, 1170 (2003).
8. Y. Banda et al., *Genetics* **200**, 1285 (2015).
9. C. E. Powe et al., *N. Engl. J. Med.* **369**, 1991 (2013).
10. D. Roberts, *Fatal Invention: How Science, Politics, and Big Business Re-Created Race in the Twenty-First Century* (The New Press, New York, 2012).
11. D. Serre, S. Pääbo, *Genome Res.* **14**, 1679 (2004).
12. P. C. Ng, Q. Zhao, S. Levy, R. L. Strausberg, J. C. Venter, *Clin. Pharmacol. Ther.* **84**, 306 (2008).
13. J. S. Kaufman, R. S. Cooper, *Am. J. Epidemiol.* **154**, 291 (2001).
14. T. R. Rebbeck, P. Sankar, *Cancer Epidemiol. Prev.* **14**, 2467 (2005).
15. L. M. Hunt, M. S. Megyesi, *J. Med. Ethics* **34**, 495 (2008).
16. A. Smart, R. Tutton, P. Martin, G. T. H. Ellison, R. Ashcroft, *Soc. Stud. Sci.* **38**, 407 (2008).
17. G. Lettre et al., *PLOS Genet.* **7**, e1001300 (2011).
18. K. Bremer, H. E. Wanntorp, *Syst. Biol.* **28**, 624 (1979).
19. C. Stewart, M. S. Pepper, *Genet. Med.* (2015).
20. G. Coop et al., *New York Times*, 8 August 2014, p. BR6.
21. T. Dobzhansky, *Mankind Evolving: The Evolution of the Human Species* (Yale Univ. Press, New Haven, CT, 1962).
22. S. M. Fullerton, J. H. Yu, J. Crouch, K. Fryer-Edwards, W. Burke, *Hum. Genet.* **127**, 563 (2010).
23. L. Braun, E. Hammonds, *Soc. Sci. Med.* **67**, 1580 (2008).
24. W. C. Byrd, M. W. Hughey, *Ann. Am. Acad. Pol. Soc. Sci.* **661**, 1 (2015).

ACKNOWLEDGMENTS

All authors contributed equally to the conceptualization of this paper. M.Y. wrote the first draft and revised it based on comments from coauthors. Special thanks to E. Arana for research assistance.

10.1126/science.aac4951