

# Changing the paradigm from 'race' to human genome variation

Charmaine D M Royal & Georgia M Dunston

**Knowledge from the Human Genome Project and research on human genome variation increasingly challenges the applicability of the term 'race' to human population groups, raising questions about the validity of inferences made about 'race' in the biomedical and scientific literature. Despite the acknowledged contradictions in contemporary science, population-based genetic variation is continually used to explain differences in health between 'racial' and 'ethnic' groups. In this commentary we posit that resolution of apparent paradoxes in relating biology to 'race' and genetics requires thinking 'outside of the box'.**

## Introduction to the state of the science

Knowledge gained from the Human Genome Project and research on human genome variation is forcing a paradigm shift in thinking about the construct of 'race'<sup>1-8</sup>, much like the process described by Thomas Kuhn in his renowned book, *The Structure of Scientific Revolutions*<sup>9</sup>. Kuhn describes the paradigm shift in science as occurring when anomalous, scientific results cannot be explained by inadequate methods. With an accumulation of such anomalies, scientists must begin to consider that the paradigm or model of reality under which the hypotheses are tested has shifted and is no longer valid. Today, scientists are faced with this situation in genomics, where existing biological models or paradigms of 'racial' and 'ethnic' categorizations cannot accommodate the uniqueness of the individual and universality of humankind that is evident in new knowledge emerging from human genome sequence variation research and molecular anthropological research. The paradigms of human identity based on 'races' as biological constructs are being questioned in light of the preponderance of data on human genome sequence variation<sup>10-13</sup> and reflect the need for a new

explanatory framework and vision of humankind with different fundamental assumptions about biological groups that can accommodate new knowledge from a new generation of research.

Discourse on the validity of 'racial' categorization in humans is certainly not new and will perhaps continue for generations to come, taking on various forms as new scientific and nonscientific knowledge emerges. Shifts have occurred over time from a purely anthropological or biological debate<sup>14-21</sup> to conversations about numerous psychosocial, societal, ethical and legal ramifications<sup>22-25</sup>, indicative of the undeniable applicability of the topic of 'race' to virtually every aspect of human existence.

This commentary describes the intellectual climate under which new information from human genome research is introduced into twenty-first-century biomedical science and society, new information that forces a more integrative construct of human biology and disease. The discordance between 'race' and human genome variation sets the stage for an analysis of the state of the science on human genome variation and 'race' and the relationship between genome variation and population differences in health and disease. The paper also provides a brief background for, and overview of, this Supplement to *Nature Genetics*.

## Genetics research and health disparities

Recent studies by leading human genome researchers report differences between African

and non-African population groups in the structure of sequence variation in the human genome<sup>26,28</sup>, rekindling in the scientific literature, as well as in the public media, old controversies over the biological relevance of 'race' in medicine. Human genome-based knowledge challenges science and society to address questions on the validity of 'racial' or 'ethnic' categories for biomedical and genetics research. It also raises questions on the public health importance of human genome variation within and between different racial or ethnic groups, thus making it relevant to the categorization of human identity in health disparities research, training and community partnership.

Research focusing on health disparities with respect to common complex diseases (such as diabetes, cardiovascular disease, and some cancers) in 'racial' or 'ethnic' groups has highlighted almost exclusively social, cultural, environmental and economic causal factors while disregarding potential genetic factors. This may be attributed to several causes, ranging from a general perception that genetics plays a minimal role, to the limitations of technology prior to the human genome project in studying the genetics of common complex diseases. The introduction of human genome technology into investigations of health disparities is controversial and demands critical examination<sup>28,29</sup>.

Human genome knowledge has been likened to a "double-edged sword"<sup>28,29</sup>, with power to exacerbate health disparities if (i) its benefits are realized only by the most affluent

National Human Genome Center, College of Medicine, Howard University, 2216 6<sup>th</sup> Street NW, Suite 207, Washington, DC 20059, USA. Correspondence should be addressed to C.D.M.R. ([croyal@howard.edu](mailto:croyal@howard.edu)).

Published online 26 October 2004;  
doi:10.1038/ng1454

members of society; (ii) its research is carried out and applied mainly toward the medical treatment of rare diseases; (iii) its message is distorted into impressions of group inferiority; (iv) its information is used to discriminate and stigmatize; and (v) its power is used to further the image of a single physical ideal. Conversely, and in line with the position of the National Human Genome Center (NHGC), this same knowledge can be effectively used to eliminate health disparities if (i) its applications are focused on common complex diseases in the least healthy groups in society; (ii) its study provides valuable insights into the causes of health disparities; (iii) its benefits are shared with vulnerable population groups; and (iv) its message is understood as valuing human variation as an instrument of self-discovery<sup>30</sup>.

### Translational genomics at the NHGC

With a mission to “explore the science of and teach the knowledge about DNA sequence variation and its interaction with the environment in the causality, prevention, and treatment of diseases common in African Americans and other African Diaspora populations”, the NHGC at Howard University was formally announced on 1 May 2001 and dedicated to the engagement of African Americans and other people of African ancestry into the mainstream of human genome research. As the only research center of its kind in a historically black academic center, the NHGC is thought to be crucial to broadening the base and active participation of African Americans and other grossly under-represented African Diaspora populations in the human genome arena. With a structure that includes genetic epidemiology, molecular genetics, bioethics, statistical genetics and bioinformatics, the NHGC was formed to address the biomedical, ethical, legal and social issues raised by the wealth of knowledge unleashed by the sequencing of the human genome.

As previously indicated, much of the current literature on genetics and health disparities emphasizes the potential dangers of connecting genetics with disparities, and relatively little research has been directed towards the potential of genomics to further understand health disparities in ways that can accomplish the US public health objectives of Healthy People 2010: a long and healthy life for all and the elimination of health disparities<sup>28,29</sup>. Conditions are prime for the application of knowledge gained from research on the structure of DNA sequence variation in African and African Diaspora populations to probe the influence of gene-environment interactions in

race- and ethnicity-based health disparities. With plans underway for the Translational Genomics Research in the African Diaspora initiative, the NHGC is positioned to lead the US and the global community with a large-scale, interdisciplinary project for human genome research in the African Diaspora. Translational Genomics Research in the African Diaspora will be a population-based resource for translational genomics in clinical research, which capitalizes on the evolutionary and migration history of Africans and the African Diaspora, and a resource for dissecting the contributions of gene-environment interactions (environment broadly defined to include psychosocial, cultural and other subjective factors) to disease susceptibility and response to medicines.

The relevance of the topic ‘genetics and race’ to the mission of the NHGC and to improved understanding of the relationships among gene-environment interactions, complex traits and health disparities between racial or ethnic groups cannot be overstated. Critics challenge the NHGC research focus on African Americans and other African Diaspora populations. Assertions that the center perpetrates race-based science and medicine have resulted from different perspectives on population-based genetic studies. The NHGC posits that the term ‘race’, as applied to humans, is incorrectly used. Traditional ‘racial’ designations in humans are not bounded, discrete categories but are fluid, socially defined constructs that have some poorly understood correlations with various biological elements and health outcomes. It is our intent that the work of the center will increase understanding of the complex interaction of genes and environment as well as cultural and other psychosocial factors that contribute to common complex diseases.

### Rationale for the NHGC meeting

In an attempt to advance the dialog among persons from various academic disciplines, professions, social strata and racial or ethnic groups, the NHGC launched a series of meetings on the conceptualization of genetic variation as ‘race’. The series is intended to bring together a diverse group of individuals over time, charged with confronting the perplexing issues from various vantage points through open and scholarly dialogue and with generating tangible outcomes.

The scientific focus of the inaugural meeting in the *Human Genome Variation and “Race”* series is by no means arbitrary. A substantial portion of the ongoing dialog on this issue has been devoted to the glaring medical

and societal implications, often glossing over the ambiguous science that underlies many of these implications. Until sound conclusions of the science are clearly communicated, society as a whole will be severely limited in its capacity to effectively address any of the ramifications. Undoubtedly, such clarity will require new approaches to methodology, training, policy and priorities. The desire of the NHGC is that the discourse initiated at the inaugural meeting and continued in this journal issue will enlighten the community about the scientific aspects of the interface between race and genetics.

### Overview of this Supplement

The papers in this Supplement are Commentaries and Perspectives from selected noted scientists and scholars in the fields of biology, human genetics, anthropology, epidemiology and bioethics, all of whom were invited presenters at the Howard University *Human Genome Variation and “Race”* meeting. This Supplement builds on and extends the discourse on the state of the science on genetics, race and health.

Mildred Cho and Pamela Sankar set the stage by highlighting the connection between genetics research and the ethical, legal and social implications of the inevitable scientific outcomes. They argue for the involvement of genetics researchers not only in the generation of knowledge about human genome in general and genetic variation research in particular, but also in the use and application of that knowledge, especially nonmedical uses such as forensic genetics.

Francis Collins, director of the National Human Genome Research Institute at the US National Institutes of Health, provides a general overview of, and model for, contextualizing existing knowledge about the interactions among race, ethnicity, genetics and health. He also proposes an agenda for additional research that is needed to advance understanding and application of these interactions and describes related efforts supported or led by the National Human Genome Research Institute.

The paper by Shomarka Keita and colleagues at the NHGC uses a historical framework in putting forward the NHGC’s position on the meaning and application of the term ‘race’. The primary assertion by these authors is that biological variation in modern humans does not structure into phylogenetic subspecies (‘races’). In addition, they point out that the controversies engendered by the term ‘race’ result primarily from problems with semantics due to inconsistency in the use and definition of the word.

In addressing the issue of whether populations cluster according to the popular concept of race, Sarah Tishkoff and Kenneth Kidd show that racial classifications do not adequately describe the distribution of genetic variation in humans. While acknowledging the clustering of populations in broad geographic regions, they contend that the broad global pattern is indicative of genetic drift associated with the African origin, followed by expansion out of Africa and across the rest of the globe. They further suggest that biomedical studies can benefit from knowledge of individual ancestry, as various factors may lead to geographical restriction of disease-associated genes.

The perspective by Lynn Jorde and Stephen Wooding emphasizes the geographic configuration of genetic variation in line with historical patterns of gene flow and genetic drift. The authors show that the distribution of genetic variation across populations is continuous and overlapping, and that observed correlations with some traditional concepts of race are limited. They also provide a general overview of patterns of human variation at the population and individual levels. They caution that although ancestry (or 'race') may prove useful in biomedicine, more accurate and beneficial information may be obtained through direct assessment of disease-related genetic variation.

Sarah Tate and David Goldstein examine the potential for pharmacogenetics to exacerbate disparities in both health and health care if measures are not instituted to ensure that the development and dispensing of medicines are inclusive. Accordingly, they call for pharmacogenetic research processes that take into account the range of ethnic and genetic diversity within and between human populations, as well as for increased participation of healthy volunteers in such studies.

Charles Rotimi presents an insider's view of the present and potential challenges related to the retention and use of racial, ethnic or population identifiers in large-scale genomic projects, such as the International Human Haplotype Map (HapMap) project. He reiterates the general consensus that racial classifications are imprecise and fluid, often correlating spuriously with genetic variation across populations. Consequently, he advocates more careful consideration of the scientific, clinical, social and ethical ramifications inherent in the design and implementation of 'race-based' population studies and the development of 'race-based' pharmacogenomic interventions.

In their assessment of scientific data on human genotypic and phenotypic variation generated over the last 35 years, Joanna Mountain and Neil Risch found that despite technological advances resulting in a large

volume of new data, progress has been relatively slow towards elucidation of the genetic basis for within- or between-group variation, particularly for complex traits and common diseases. They attribute this to intrinsic difficulty in teasing out the associated genetic influences, as well as to the important role of nongenetic factors. They conclude that given the ongoing challenges in understanding on the role of genes in between-group variation, generalizations regarding genetic contributions to observed differences are unwarranted and may exacerbate group disparities.

Esteban Parra, Rick Kittles and Mark Shriver present results of their study designed to evaluate correlations between skin pigmentation and ancestry. They observed substantial but variable strengths of correlation between pigmentation and ancestry in each of the five populations studied, attributing their observations to varying degrees of admixture stratification among populations or differences in the levels of pigmentation between the parental populations and the number of genes involved. They recommend caution when using pigmentation as a 'marker' of ancestry or when extrapolating the results from one population to other admixed populations.

### Conclusion

Reflecting on this compilation of articles from the distinguished group of scientists invited to contribute to this special issue of *Nature Genetics*, it is evident that much effort has already been expended in attempting to achieve clarity on the complex relationships among race, genetics and health. Based on information presented here, there seems to be consensus that 'race', whether imposed or self-identified, is a weak surrogate for various genetic and nongenetic factors in correlations with health status. We are at the beginning of a new era in molecular medicine. It remains to be determined how increasing knowledge of genetic variation in populations will change prevailing paradigms of human health and identity. ■

### COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 10 September; accepted 23 September 2004  
Published online at <http://www.nature.com/naturegenetics/>

- Bamshad, M., Wooding, S., Salisbury, B.A. & Stephens, J.C. Deconstructing the relationship between genetics and race. *Nat. Rev. Genet.* **5**, 598–609 (2004).
- Burchard, E.G. *et al.* The importance of race and ethnic background in biomedical research and clinical practice. *N. Engl. J. Med.* **248**, 117–1075 (2003).
- Cooper, R.S., Kaufman, J.S. & Ward, R. Race and

- genomics. *N. Engl. J. Med.* **348**, 1166–1170 (2003).
- Foster, M.W. & Sharp, R.R. Race, ethnicity, and genomics: Social classifications as proxies of biological heterogeneity. *Genome Res.* **12**, 844–850 (2002).
- Frank, R. The misuse of biology in demographic research on racial/ethnic differences: a reply to Van Den Oord and Rowe. *Demography* **38**, 563–567 (2001).
- Goodman, A.H. Why genes don't count for racial differences in health. *Am. J. Public Health* **90**, 1699–1702 (2000).
- Phimister, E.G. Medicine and the racial divide. *N. Engl. J. Med.* **348**, 1081–1082 (2003).
- Kittles, R.A. & Weiss, K.M. Race, ancestry, and genes: implications for defining disease risk. *Annu. Rev. Genomics Hum. Genet.* **4**, 33–67 (2003).
- Kuhn, T. *The Structure of Scientific Revolutions* (University of Chicago Press, 1962).
- Barbujani, G., Magagni, A., Minch, E. & Cavalli-Sforza, L. An apportionment of human DNA diversity. *Proc. Natl. Acad. Sci. USA* **94**, 4516–4519 (1997).
- Rosenberg, N. *et al.* Genetic structure of human populations. *Science* **298**, 2381–2385 (2002).
- Jorde, L.B. *et al.* The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *Am. J. Hum. Genet.* **66**, 979–988 (2000).
- Bamshad, M.J. *et al.* Human population genetic structure and inference of group membership. *Am. J. Hum. Genet.* **72**, 578–589 (2003).
- Livingstone, F.B. On the non-existence of Human Races. *Curr. Anthropol.* **3**, 279–281 (1958).
- Wiercinski, A. & Bieliński, T. The racial analysis of human populations in relation to their ethnogenesis. *Curr. Anthropol.* **3**, 29–46 (1962).
- Nash, M. Race and the ideology of race. *Curr. Anthropol.* **3**, 285–288 (1962).
- Newman, M.T. Geographic and microgeographic races. *Curr. Anthropol.* **4**, 189–207 (1963).
- Gates, R.R. Racial genetics: a new branch of anthropology. *Curr. Anthropol.* **4**, 208–209 (1963).
- Dobzhansky, T., Montagu, A. & Coon, C.S. Two views of Coon's origin of races with comments by Coon and replies. *Curr. Anthropol.* **4**, 360–367 (1963).
- Brace, C.L. On the race concept. *Curr. Anthropol.* **5**, 313–320 (1964).
- Edmonson, M.S. A measurement of relative racial difference. *Curr. Anthropol.* **6**, 167–180, 181–198 (1965).
- Cooper, R. A note on the biologic concept of race and its application in epidemiologic research. *Am. Heart J.* **108**, 715–723 (1984).
- Smedley, A. *Race in North America: Origin and Evolution of a Worldview* 2<sup>nd</sup> edn. (Westview, Boulder, Colorado, 1999).
- Duster, T. *Backdoor to Eugenics* 2<sup>nd</sup> edn. (Routledge, New York, New York, 2003).
- Ellison, G.T.H. & Jones, I.R. Social identities and the 'new genetics': Scientific and social consequences. *Critical Public Health* **12**, 265–282 (2002).
- Tishkoff, S.A. & Verrelli, B.C. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annu. Rev. Genomics Hum. Genet.* **4**, 293–340 (2003).
- Tishkoff, S.A. & Williams, S.M. Genetic analysis of African populations: human evolution and complex disease. *Nat. Rev. Genet.* **3**, 611–621 (2002).
- Citrin, T. & Modell, S. Genomics and public health: Ethical, legal, and social issues. in *Genomics and Population Health: United States 2003* (eds. Gwinn, M., Bedrosian, S., Ottman, D. and Khoury, M.) (Office of Genomics and Disease Prevention, Center for Disease Control, 2003).
- Lee, S.S., Mountain, J. & Koenig, B.A. The meanings of "race" in the new genomics: implications for health disparities research. *Yale J. Health Policy Law Ethics* **1**, 33–75 (2001).
- Dunston, G.M. & Royal, C.D.M. The human genome: implications for the health of African Americans. in *The Praeger Handbook of Black American Health: Policies and Issues Behind Disparities in Health* vol. 2, 2<sup>nd</sup> edn. (ed. Livingston, I.L.) 757–775 (Greenwood Publishing Group, 2004).