Understanding and Using Human Genetic Variation Knowledge in the Design and Conduct of Biomedical Research.

By Charles N. Rotimi, PhD

The successful completion of the sequencing of the human genome has taken us into a new world, a world in which we are challenged to look at ourselves in ways we never before thought possible. Knowledge gained from genomic science promises to have a profound impact on our understanding of issues surrounding self and group identity. Also, genome-based data will inform our understanding of ethnicity and race and, by implication, racism. We can now query our genome—our entire genetic make-up—to answer important questions: Where are we from? How are we related? And why does disease burden vary by family, ethnic and ancestral groups? This new knowledge has the potential to redefine the ways we conduct biomedical research including clinical trials and the development and administration of new medications at the individual and group levels.

Human Genetic Variation and Group Identity

Following the complete sequencing of the human genome, scientists are faced with the major challenge of documenting, describing and understanding the non-random pattern of human genetic variation and its link to disease risks in different populations. Findings from the huge amount of genetic data generated so far in various groups show clearly that most (> 90%) of the observed genetic variations occur within rather than between groups (< 10%; estimates are as low as 3%). [1]

A critical question that must be answered by geneticists and others is: Does the level of genetic variation observed between groups rise to the level of sub-speciation? In other words, do human groups structure biologically into subspecies?

Although population genetic data overwhelmingly show that it is difficult to consistently identify all members of a group without generating "outliers," scientists have been unable to move beyond racial categorization in science, medicine and society. Partially responsible for our continued obsession with race is the fact that, although we do not have distinct biological types of "races," we do have differences in the frequencies of genetic markers across human ancestral groups. These differences, which for the most part describe continental populations (geographical distance), are believed to harbor the answers to why some individuals and groups may be more susceptible or resistant to diseases and may also hold the key to understanding why different human groups respond differently to some medications.

The well documented small genetic differences between continental populations (i.e., about 10% of the 0.1% of our DNA that varies among individuals) have been used by some scientists to justify the social construct of "race." [2] Hence, it is important to point out that recent data support the notion that the human evolutionary tree is a virtual continuum of genetic variations and that because most alleles (different forms of a gene) are widespread, genetic differences among human populations are the result of gradations in allele frequencies rather than distinctive diagnostic genotypes.[3]

An understanding and appreciation of this knowledge prompted scientists at the National Human Genome Center (NHGC) at Howard University to convene a national meeting on "Human Genetic Variation and 'Race': The State of the Science." Experts in history, anthropology, genetics, epidemiology and medicine were brought together to discuss the relationship between observed genetic variation and 'race'. This very successful meeting resulted in the publication of several scholarly papers in a 2004 special supplement to the journal *Nature Genetics*. Following the meeting, scientists at the NHGC at Howard University developed a position paper on race and genetics. Table 1 summarizes the observations and recommendations included in the position paper.

"The solution to the challenges presented by race is not to ignore group differences in disease patterns but to find innovative ways to study the complex interplay between genetic and environmental factors in the etiology of various health outcomes."

Health Disparity and Genomic Medicine

Despite our inability to classify humans into distinct groups by genes, race and racism have strong cultural, political and economic significance with real biological consequences. As used in medical research, the term "race" has "real" meaning and racial categories, however flawed, are acceptable descriptive labels. Given the practical implication of race for the health of individuals and of population groups, the question becomes how to develop strategies to study its impact.

Before embarking on this, however, we should consider whether we have a better descriptive term for the social phenomenon called "race." Ethnicity, a social construct with some similarities to race, incorporates multiple variables including genetics, economic, social, religious, and linguistic background, and dietary habits to identify individuals as belonging to a group or population. While it is imprecise, "ethnicity" captures better than "race" more of what may indeed be responsible for observed health disparity between population groups and allows for potential differences in gene frequencies and for dynamism—a hallmark of human evolutionary history.

The solution to the challenges presented by race is not to ignore group differences in disease patterns but to find innovative ways to study the complex interplay between genetic and environmental factors in the etiology of various health outcomes. The definition of environmental factors should be expanded to incorporate non-traditional epidemiologic variables including some measure of self-identity and its impact on health risks. The new challenge—understanding the etiology of common complex

traits (e.g., diabetes and hypertension)—demands interdisciplinary approaches and the development of innovative ways of studying the rules of complex systems without taking them apart.

While we will always find group differences that may lend themselves to simplistic stereotypes, we must resist the old temptation of explaining group differences as innate, especially in the context of the historic experiences of the peoples of the Western Hemisphere including the United States of America. As articulated by Richard Lewontin, "[R]ace, ethnicity, and social class are so confounded, and the reality of social class so firmly denied, that it is easy to lose sight of the general setting of class conflict out of which biological determinism arose. Biological determinism, both in its literary and scientific forms, is part of the legitimating ideology of our society, the solution offered to our deepest social mystery, the analgesic for our most recurrent social pain." [4] Thus, it is imperative that we do a better job of documenting the social and cultural ills responsible for health disparity at national and global levels, before invoking genetic explanations for observed group differences in disease distributions. In the words of Charles Darwin, quoted on the title page of *The Mismeasure of Man*, *"If the misery of our poor be caused not by the laws of nature, but by our institutions, great is our sin."* [4]

Pharmacogenomics: genes, drugs and group identity

As the human genome project and subsequent spin-offs (e.g., HapMap) help us to develop better drugs, it is important that we ask ourselves pointedly: "Will tomorrow's medicines work for everyone?" To answer this question we must develop strategies to catalog and understand the influence of genetic and non-genetic factors on individual and group responses to treatments. Research design strategies must also allow scientists to adequately represent the spectrum of genetic variation across multiple human populations. *It is, therefore, important to include multiple ethnic groups in clinical trials and specifically, pharmacogenomics research projects not because these are distinct biological groups but rather because there are subtle differences in allele frequencies between groups that may be important in how members of these groups respond to drugs at the individual level.* These potential genetic differences need to be understood both at the individual and group levels. Although non-genetic factors (e.g., age, organ function and nature of the disease) can influence an individual's drug reaction, evidence demonstrates that genetics account for most of the observed variability in individual drug disposition and effects.[5,6] In addition, genetic factors influencing drug response are more stable throughout a person's lifetime as compared to non-genetic factors[5] and thus, are more likely to provide better long-term information on pharmacokinetics and pharmacodynamics of drugs.

It is instructive to evaluate the evidence in favor of how frequency differences in some known pharmacogenetic variants may influence individual and ethnic group response to drugs. The first example is the N-acetyltransferase 2 (NAT2); this enzyme is involved in the detoxification of many carcinogens and the metabolism of many commonly used drugs. It has been known for sometime that genetic variation of NAT2 results in two phenotypes, slow and rapid acetylators.[7,8] Interestingly, there are significant group differences in the distribution of the slow acetylator phenotype ranging in frequency from about 14% in East Asians, 34% in African Americans and to a high of 54% among European Americans. Given this data it is reasonable to suggest that these frequency differences may have significant impact on how these ethnic or ancestral groups may experience the toxic effect of drugs and their susceptibility to environmental carcinogens.

A second example is variation in the *CCR5* gene—a receptor used by human immunodeficiency virus (HIV) to enter cells. Current data on a variant of this gene, *CCR5-delta32*, which offers some protection against HIV infection and progression, varies considerably between ethnic groups.[9] About 25% of European Americans have this allele while it is virtually absent in other ethnic groups.[10] If this variant holds as a major protector against HIV infection, it is also reasonable to suggest that it may help explain differences in ethnic susceptibility to HIV infection.

As it becomes more feasible to use high-throughput technologies to genotype individuals, the conduct of future clinical trials may be based on patient groups characterized by genetic variation information instead of the current imprecise classifications such as "race", ethnicity, or similar groupings of individuals. As we achieve this level of individual understanding of genetic variation within the historic context of how these persons live their lives, we should experience a reduction in the frequency of adverse reactions to drugs. Furthermore, understanding the detailed structure of human genetic variation may help to deconstruct the notion of "race" and other imprecise group definitions as currently applied in biomedical research. ●

Acknowledgement:

The National Human Genome Center position statement was drafted by the following scientists: Royal CMD, Keita SOY, Kittles RA, Dunston GMD and Rotimi CN.

EDICT
Eliminating Disparities in Clinical Trials

References

1. Jorde LB and Wooding SP. Genetic variation, classification and 'race'. Nat Genet. 2004 Nov;36(11 Suppl):S28-33.
2. Rotimi CN. Are medical and nonmedical uses of large-scale genomic markers conflating genetics and 'race'? Nat Genet. 2004; 36(11 Suppl):S43-7.
3. Rosenberg NA, et al. Genetic structure of human populations. Science. 2002; 298(5602):2381-5.

4. Lewontin R. It Ain't Necessarily So – The dream of the Human Genome and other Illusions pp18-19. Second Edition. New York Review Books, New York.
5. Evans WE and McLeod HL. Pharmacogenomics--drug disposition, drug targets, and side effects. N Engl J Med. 2003;348 (6):538-49.
6. Kalow W et al. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. Pharmacogenetics. 1998 Aug;8(4):283-9.
7. Burchard EG et al. The importance of race and ethnic background in biomedical research and clinical practice. N Engl J Med. 2003 Mar 20;348(12):1170-5.
8. Yu MC et al. Acetylator phenotype, aminobiphenyl-hemoglobin adduct levels, and bladder cancer risk in white, black, and Asian men in Los Angeles, California.
J Natl Cancer Inst. 1994 May 4;86(9):712-6.
9. Dean M, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. Science. 1996 Sep 27;273(5283):1856-62.
10. Stephens JC 1998 et al. Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. Am J Hum Genet. 1998; 62(6):1507-15.

Table 1

Observations and recommendations regarding race and genetics by the National Human Genome Center of Howard University
1. When the human species is viewed as a whole, underlying genetic variation and expressed physical traits exhibit gradients of differentiation, not discrete units. Therefore, modern extant humans do not fracture into races (subspecies) based on the modern phylogenetic criteria of molecular systematics.
2. The biological "boundaries" between any human divisions (groups, populations, nationalities) are circumstantial and largely dependent on what traits are chosen for emphasis.
3. The demographic units of human societies (and of the U.S. census) are the products of social or political rules, not the forces of biological evolution. The names and characteristics of demographic groups can change and have changed over time.
4. Group differences in health parameters are not encoded in the human genome as part of an evolutionary pattern of divergence. Thus, differences in health or disease cannot be treated as causally related to ethnoancestral groups.
5. Genotype-environment interactions are more important in explaining group differences in health than genotype, environment, or a factor called "race".
6. The non-existence of human races (subspecies) does not mean the non-existence of racism. Racism is the structured systematic oppression against individuals and groups defined based on physical traits that reflect an extremely limited fraction of the human genome. Racism must be addressed.
7. Individuals cannot be treated as representative for all those who physically resemble them, or have some of the same ethnohistorical ancestry. Ancestries of individuals and groups should be ascertained in order to evaluate differential expression of genetic effects.