

# Confusions About Human Races

By R.C. Lewontin

*R.C. Lewontin, Alexander Agassiz Professor Emeritus of Zoology at Harvard University, has written a number of books and articles on evolution and human variation, including Biology as Ideology: The Doctrine of DNA and The Triple Helix: Gene, Organism, and Environment*

Over the last thirty five years a major change has taken place in our biological understanding of the concept of human “race,” largely as a consequence of an immense increase in our knowledge of human genetics. As a biological rather than a social construct, “race” has ceased to be seen as a fundamental reality characterizing the human species. Nevertheless, there appear from time to time claims that racial categories represent not arbitrary socially and historically defined groups but objective biological divisions based on genetic differences. The most recent widely noticed rebirth of such claims is an essay by Armand Marie Leroi on the Op-Ed page of *The New York Times* (March 14, 2005), an essay that illustrates both the classical confusions about the reality of racial categories and the more recent erroneous conclusions about the relevance of such racial identifications for medical practice.

There are four facts about human variation upon which there is universal agreement. First, the human species as a whole has immense genetic variation from individual to individual. Any two unrelated human beings differ by about 3 million distinct DNA variants.

Second, by far the largest amount of that variation, about 85%, is among individuals within local national or linguistic populations, within the French, within the Kikuyu, within the Japanese. There is diversity from population to population in how much genetic variation each contains, depending upon how much immigration into the population has occurred from a variety of other groups and also on the size of the population. The United States, with a very large population whose ancestors came from all over the earth including the original inhabitants of the New World, is genetically very variable whereas small populations of local Amazonian tribes are less genetically variable, although they are by no means genetically uniform. Despite the differences in amount of genetic variation within local populations, the finding that on the average 85% of all human genetic variation is within local populations has been a remarkably consistent result of independent studies carried out over twenty-five years using data from both proteins and DNA.

Of the remaining 15% of human variation, between a quarter and a half is between local populations within classically defined human “races,” between the French and the Ukrainians,

between the Kikuyu and the Ewe, between the Japanese and the Koreans. The remaining variation, about 6% to 10% of the total human variation is between the classically defined geographical races that we think of in an everyday sense as identified by skin color, hair form, and nose shape. This imprecision in assigning the proportion of variation assigned to differences among population within "races" as compared to variation among "races," arises precisely because there is no objective way to assign the various human populations to clear-cut races. Into which "race" do the Hindi and Urdu speakers of the Indian sub-continent fall? Should they be grouped with Europeans or with Asians or should a separate race be assigned to them? Are the Lapps of Finland and the Hazari of Afghanistan really Europeans or Asians? What about Indonesians and Melanesians? Different biologists have made different assignments and the number of "races" assigned by anthropologists and geneticists has varied from 3 to 30.

Third, a small number of genetic traits, such as skin color, hair form, nose shape (traits for which the genes have not actually been identified) and a relatively few proteins like the Rh blood type, vary together so that many populations with very dark skin color will also have dark tightly curled hair, broad noses and a high frequency of the Rh blood type R0. Those who, like Leroi, argue for the objective reality of racial divisions claim that when such covariation is taken into account, clear-cut racial divisions will appear and that these divisions will correspond largely to the classical division of the world into Whites, Blacks, Yellows, Reds and Browns. It is indeed possible to combine the information from covarying traits into weighted averages that take account of the traits' covariation (technically known as "principal components" of variation). When this has been done, however, the results have not borne out the claims for racial divisions. The geographical maps of principal component values constructed by Cavalli, Menozzi and Piazza in their famous *The History and Geography of Human Genes* show continuous variation over the whole world with no sharp boundaries and with no greater similarity occurring between Western and Eastern Europeans than between Europeans and Africans! Thus, the classically defined races do not appear from an unprejudiced description of human variation. Only the Australian Aborigines appear as a unique group.

A clustering of populations that does correspond to classical continental "races" can be achieved by using a special class of non-functional DNA, microsatellites. By selecting among microsatellites, it is possible to find a set that will cluster together African populations, European populations, and Asian populations, etc. These selected microsatellite DNA markers are not typical of genes, however, but have been chosen precisely because they are "maximally informative" about group differences. Thus, they tell us what we already knew about the differences between populations of the classical "races" from skin color, face shape, and hair

form. They have the added advantage of allowing us to make good estimates of the amount of intermixture that has occurred between populations as a result of migrations and conquests.

The every-day socially defined geographical races do identify groups of populations that are somewhat more closely similar to each other genetically. Most important from the standpoint of the biological meaning of these racial categories, however, most human genetic variation does not show such "race" clustering. For the vast majority of human genetic variations, classical racial categories as defined by a combination of geography, skin color, nose and hair shape, an occasional blood type or selected microsatellites make no useful prediction of genetic differences. This failure of the clustering of local populations into biologically meaningful "races" based on a few clear genetic differences is not confined to the human species. Zoologists long ago gave up the category of "race" for dividing up groups of animal populations within a species, because so many of these races turned out to be based on only one or two genes so that two animals born in the same litter could belong to different "races."

In his article, Leroi is inconsistent and shifting in his notion of race. Sometimes it corresponds to the classical social definitions of major races, but elsewhere he makes "race" coincident with a small local group such as the Negritos or Inuit. In this shifting concept of "race" he goes back to the varying use of the term in the 19th century. Then people spoke of the "Scots race," "the Irish race" and the "race of Englishmen." Indeed "race" could stand for a family group defined by male inheritance, as in the description of the last male in a family line as "the last of his race." This inconsistent usage arises from the fact that there is no clear criterion of how much difference between groups of genetically related individuals should correspond to the category "race." If it had turned out that groups of related populations were clearly different in the great majority of their genes from other groups, then racial categories would be clear and unambiguous and they would have great predictive power for as yet unstudied characters. But that is not the way it has turned out, at least for the human species.

The fourth and last fact about genetic differences between groups is that these differences are in the process of breaking down because of the very large amount of migration and intergroup mating that was always true episodically in the history of the human species but is now more widespread than ever. The result is that individuals identified by themselves or others as belonging to one "race," based on the small number of visible characters used in classical race definitions, are likely to have ancestry that is a mixture of these groups, a fact that has considerable significance for the medical uses of race identification.

A common claim, repeated by Leroi, is that racial categories are of considerable medical use, especially in diagnostic testing because some genetic disorders are very common in ancestral racial populations. For example sickle cell anemia is common among West Africans, who were brought as slaves to the New World, and Tay-Sachs disease is common among Ashkenazi Jews. So, it is argued, racial information can be a useful diagnostic indicator. Certainly classical “race” contains some medically relevant information in some cases, as for example “white” as opposed to “African American” if the contrast is between Finland and West Africa, but not if it is a contrast between a “white” Mediterranean and an “Asian” Indian. There is a confusion here between race and ancestry. Sickle cell anemia is in high frequency not only in West Africans but also in some “white” Middle Eastern and Indian populations. Moreover, a person with, say, one African great-grandparent, but who is identified by herself and others as “white” has a one in eight chance of inheriting a sickle-cell mutation carried by that ancestor. There are, in addition, a number of other simply inherited hemoglobin abnormalities, the thalassemys, that are in high frequency in some places in the Mediterranean (Sardinia), Arabia and southeast Asia. The highest frequency known for a thalassemy (80%) is in Nepal, but it is rare in most of Asia. The categorization of individuals simply as “white” or “Afro-American” or “Asian” will result in a failure to test for such abnormal hemoglobins because these abnormalities do not characterize the identified “race” of the patient. Even group identities below the level of the conventional races are misleading. Two of my incontrovertibly WASP grandchildren have a single Ashkenazi Jewish great-grandparent and so have a one in eight chance of inheriting a Tay-Sachs abnormality carried by that ancestor. For purposes of medical testing we do not want to know whether a person is “Hispanic” but rather whether that person’s family came from a Caribbean country such as Cuba, that had a large influx of West African slaves, or one in which there was a great deal of intermixture with native American tribes as in Chile and Mexico, or one in which there was only a negligible population of non-Europeans. Racial identification simply does not do the work needed. What we ought to ask on medical questionnaires is not racial identification, but ancestry. “Do you know of any ancestors who were (Ashkenazi Jews, or from West Africa, from certain regions of the Mediterranean, from Japan)?” Once again, racial categorization is a bad predictor of biology.

There has been an interesting dialectic between the notion of human races and the use of race as a general biological category. Historically, the concept of race was imported into biology, and not only the biology of the human species, from social practice. The consciousness that human beings come in distinct varieties led, in the history of biology, to the construction of “race” as a subgrouping within species. For a long time the category “race” was a standard taxonomic level. But the use of “race” in a general biological context then reinforced its application to humans.

After all, lots of animal and plant species are divided into races, so why not *Homo sapiens*? Yet the classification of animal and plant species into named races was at all times an ill-defined and idiosyncratic practice. There was no clear criterion of what constituted a race of animals or plants that could be applied over species in general. The growing realization in the middle of the twentieth century that most species had some genetic differentiation from local population to local population led finally to the abandonment in biology of any hope that a uniform criterion of race could be constructed. Yet biologists were loathe to abandon the idea of race entirely. In an attempt to hold on to the concept while make it objective and generalizable, Th. Dobzhansky, the leading biologist in the study of the genetics of natural populations, introduced the “geographical race,” which he defined as any population that differed genetically in any way from any other population of the species. But as genetics developed and it became possible to characterize the genetic differences between individuals and populations it became apparent, that *every* population of every species in fact differs genetically to some degree from every other population. Thus, every population is a separate “geographic race” and it was realized that nothing was added by the racial category. The consequence of this realization was the abandonment of “race” as a biological category during the last quarter of the twentieth century, an abandonment that spread into anthropology and human biology. However, that abandonment was never complete in the case of the human species. There has been a constant pressure from social and political practice and the coincidence of racial, cultural and social class divisions reinforcing the social reality of race, to maintain “race” as a human classification. If it were admitted that the category of “race” is a purely social construct, however, it would have a weakened legitimacy. Thus, there have been repeated attempts to reassert the objective biological reality of human racial categories despite the evidence to the contrary.