nature genetics

The unexamined 'Caucasian'

Judith Hall made the point that insufficient clinical information often limits the use of genetic research (*Nat. Genet.* **33**, 440-442; 2003). Still another kind of inattention threatens the public understanding of genetic research and impairs the translation of its conclusions to other geneticists - and to clinicians, epidemiologists, anthropologists and historians - namely inadequate labeling of human populations in lieu of genotype or stated environmental factors.

On April 22nd, the Science Journal Editors' Working Group on Human Genetic Variation Research met at the Center for Bioethics at the University of Pennsylvania. Organizers Pamela Sankar and Mildred Cho suggested one of our editorials (*Nat. Genet.* **24**, 97-98; 2000) calling for equal, consistent and informative description of human populations and ethnic groups, was without effect, on the basis of samples of eight papers published before and after the editorial.

We therefore examined 60 papers published in the past year that included descriptions of geographic or ethnic groups of people, and found 22 papers that conform to both Sankar and Cho's (*Science* **298**, 1337-1338; 2002) and our recommendations. One association study matched cases and controls by age, ethnic and geographic origin: Swedish by citizenship for three generations, British by surname and citizenship, and Belgian by three generations of citizenship and Flemish self-reported ethnicity. In another paper characterizing a mutation with a founder effect in an isolated population of mixed ancestry, endogamous Cayman Islanders were compared genetically to their ancestral population of Jamaicans and these were then compared with their populations of origin in Ghana, Cameroon and Britain.

The focus of the research does matter. In epidemiology, in association studies and cases where finding differences between population groups is the aim, authors both do and need to take more care with labels and criteria describing populations. For example, in a paper examining population stratification by genomic control and by self-reported ancestry, criteria and descriptions for group membership were equally and carefully used throughout. Recommendation of best practice to other researchers was a key aim of this work, but all three papers are exemplary.

On the other hand, rare mutations of large phenotypic effect are often considered without reference to the background on which they arose. Many papers dealing with mendelian traits list only the collecting center or use geographical or ethnic labels. We favor specifying as much information as is compatible with ethical review board requirements, because although gene identification may be complete with the publication of your paper, modifier loci may be found where mutations arose in populations with different ancestral origins or are carried into a different ethnic environment. This project then leads to an understanding of the molecular and environmental context in which the gene functions.

Although the problems we identified have no single solution, we suggest three concepts to consider. First, 'ancestry' invites examination of descent, continental origin and admixture. It also invites specification of depth, for example 'African diaspora from the 15th century to the present', or 'grandparents born in Turkey'. Secondly, 'ethnicity' deals with social and cultural factors that may exert environmental influences on phenotype, migration and reproductive patterns. For example, 'self-reported Vlax Roma ethnicity'. Once ancestry and ethnicity have been described, the residual information in 'race' is the legacy of discrimination. If it is necessary to examine this, it can be described using a combination of ancestry, phenotype and ethnicity. Lastly, equality of description is needed to secure socially responsible press coverage for genetics research, but, for the purposes of portability between genetics researchers, noting the size and diversity of the groups being compared may be more useful.

Specify information about your study groups so it can be reassessed in different ways, because broad categories cannot subsequently be analyzed. Enable hypothesis generation by other researchers and avoid misrepresenting scientific ideas of human genetic variation to the media and the public. Your paper will then have more impact and influence, and its proper place in history. ■