The Environmental Genome Project: Suggestions and Concerns

The NIEHS held a Symposium on the Environmental Genome Project on 17–18 October 1997 at the National Institutes of Health in Bethesda, Maryland. The purpose of the meeting was to facilitate a free exchange of information about the Environmental Genome Project among "a diverse group of scientists working in the areas of genetics, gene–environment interactions, molecular epidemiology, and issues of genetic testing" (*1*). After attending this meeting and reading the recent reports on the Environmental Genome Project in *Science* (*2*) and *Nature* (*3*), we are concerned that several important issues of interest to the readers of *Environmental Health Perspectives* are not being adequately addressed by the NIEHS.

The goal of the Environmental Genome Project is to "facilitate identification of functionally important polymorphisms in environment response genes that may determine differences in disease risks to environmental exposures" (*1*). To this end, the NIEHS proposes to establish a repository of 1,000 anonymous DNA samples representing the population of the United States in order to catalog allelic variants in 200 genes and foster epidemiologic studies of gene–environment interactions in the etiology of human diseases. At the symposium and in the *Science* article (*2*), the potential importance of gene–environment interactions in various diseases was recognized and the potential benefits of a central repository of data on a set of critical risk-mediating genes were clearly enunciated. Two issues, however, have not been adequately addressed: sample size and public policy implications.

We have serious doubts that the Environmental Genome Project's proposed sample size of 1,000 individuals is large enough to provide stable estimates of allele frequencies in subgroups of the population. Many of the most promising candidate susceptibility genes have allelic variations that affect less than 5% of the population, and the prevalence of many of these polymorphisms differs markedly among ethnic groups. If 1% of the people in the U.S. population carry a polymorphism for a certain gene, for example, we would expect 10 individuals out of the 1,000 individuals in the study to carry that polymorphism. If those 10 are then to be subdivided into gender/racial/ethnic/subgroups, the sample size is clearly inadequate to provide precise estimates of the prevalence of the polymorphism in the subgroups. Yet precise estimates of the population prevalence are exactly the kind of information needed by epidemiologists planning studies of the role of such genes in disease etiologies. That information is often difficult to find or nonexistent in the current literature, much of which is characterized by small, nonpopulation-based studies that are difficult to generalize. It is difficult to conceive why the Environmental Genome Project would spend so much time and money only to find that 1,000 subjects were not nearly enough. Given the potential importance of this population-based data, and cognizant of the budgetary constraints, it would seem wiser to ensure the value of the data by increasing the sample size and decreasing the number of genes targeted for sequencing.

Concerning the sampling procedure, the NIEHS appears to have planned to collect samples from individuals rather than using existing archived samples. This will involve a considerable expense. An alternative, cost-saving strategy, which does not appear to have been seriously explored, would be to use the archived lymphocyte cell lines from the National Health and Nutrition Examination Survey (NHANES) III study as the source of DNA for the Environmental Genome Project. The NHANES III study population is large (over 8,000) and is population-based, a representative sample of the entire U.S. population. A further advantage of using this database is that genotype data can be linked to an enormous database of health and nutrition variables. Strategies for preserving anonymity are currently being investigated by the Centers for Disease Control and Prevention.

We are also concerned about the societal ramifications of the Environmental Genome Project. While the recent symposium did address certain ethical and policy issues such as insurance discrimination, other issues unique to the Environmental Genome Project received little attention. One of the project's goals is to improve risk assessment and regulation by government agencies through specific information about vulnerable populations. How this information is to be used in risk-based regulation in a heterogeneous society is as yet unclear. Should employers be able to transfer chemically sensitive workers to jobs with lower exposure levels rather than reducing exposure levels to a safe level for all? Could employers expose more resistant workers to higher exposure levels? In cases of alleged chemical injury, would lawyers misuse knowledge of genetic susceptibilities? The Environmental Genome Project, of course, cannot be held accountable for such misuses, but since the project aims to follow the lead of the Human Genome Project in channeling some of its resources into exploring ethical concerns, it seems logical to focus on ethical issues specific to the Environmental Genome Project before work progresses further. This may require educating courts and regulators about what can and cannot be known about risk-mediating genes in given individuals and populations.

Basic scientists have rarely had to address broad public policy issues resulting from their investigations, but genetics has increasingly become, like epidemiology, as much a tool for public health as a scientific discipline. The point on which everyone seems to agree is that improving public health is the primary goal of biomedical research. It will take careful thinking to ensure that the Environmental Genome Project serves this purpose.

Christopher A. Loffredo
Ellen K. Silbergeld
Program in Human Health and the Environment
University of Maryland School of Medicine
Baltimore, Maryland

Mark Parascandola
NIH Historical Office
National Institutes of Health
Bethesda, Maryland

References and Notes

1. NIEHS. Symposium on the Environmental Genome Project [brochure]. Research Triangle Park, NC: National Institute of Environmental Health Sciences, 1997.
2. Kaiser J. Environment institute lays plans for gene hunt. Science 278:569–570 (1997).
3. Wadman M. Genome study maps chemical sensitivity. Nature 389:774 (1997).

Response: Environmental Genome Project

The letter from Loffredo, Silbergeld, and Parascandola raises several interesting issues that I would like to address in the context of the overall concept of the Environmental Genome Project. The Environmental Genome Project is an outgrowth of a long-standing interest on the part of the NIEHS and the larger scientific community in the relationship between environmental exposure and disease and the influence of genetics upon this relationship. In their letter, Loffredo, Silbergeld, and Parascandola express concern about whether the sample size proposed for initial studies by the Environmental Genome Project is sufficient for assignment of allele frequencies, and they urge caution in dealing with the ethical, social, and legal implications of the