Correspondence

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? Can 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 longstanding 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
proposed research. These are important issues that should be addressed by the Environmental Genome Project. Issues such as these should be carefully considered and addressed during the early stage of the project. As noted at the beginning of the Environmental Genome Project symposium, the project is evolving and it is a work in progress. As with any evolving program, it will continue to be molded by new ideas, information, and technologies.

The goal of the initial phase of the Environmental Genome Project is to stimulate research in the area of polymorphism discovery. This phase of the project does not specifically seek to assign polymorphism frequency. An allele has to be present only once in the repository to be discovered, yet accurately estimating the frequency of an allele in different ethnic groups requires genotyping of large numbers of individuals. Once polymorphisms (or alleles) have been discovered, study groups can be held to consider the research required for assignment of allele frequency. While the Environmental Genome Project does not seek to assign allele frequencies, we are aware of the importance of accurate allele frequency estimates for future epidemiologic studies and the large sample sizes such estimates will require. It is important to consider whether the sample size selected by the Environmental Genome Project will provide sufficient power to discover most alleles relevant to gene–environment interactions. Clearly, sampling 500 to 1000 individuals will be adequate to identify many new polymorphisms. As pointed out during the symposium, sampling this number of individuals is adequate to identify most of the polymorphisms occurring commonly in the U.S. population. All of the newly identified polymorphisms will have the potential to be involved in gene–environment interactions, although none of them will be guaranteed to be so involved. It is also clear that various combinations of alleles may uniquely collabor ate in environmentally associated disease. These allele combinations will, of course, be present in the population at lower levels than each allele alone. As was clearly pointed out at the Environmental Genome Project symposium, studies of gene–environment or gene–gene interactions will require very large sample sizes. Planning for both the technical means of rapid genotyping and the large number of samples for future epidemiologic studies is a key component of the Environmental Genome Project.

In their letter, Loffredo, Silbergeld, and Parascandola mention the potential uselessness of the archived lymphocyte cell lines from the NHANES III study and they suggest that this collection has not been explored as a resource for DNA samples. In fact, NHANES III cell lines will be used by the Environmental Genome Project. A repository of samples is being identified by the National Human Genome Research Institute, National Institutes of Health, in partnership with the Centers for Disease Control and Prevention. More than one-half of the repository of samples will be from NHANES III.

As pointed out by Loffredo, Silbergeld, and Parascandola, another challenge to the Environmental Genome Project is in the area of its ethical, legal, and social implications. Symposium attendees discussed this topic in detail. These issues are complex and many-layered. It is highly unlikely that all the layers and nuances of the issues have been uncovered or that they will become simpler as the project evolves. To do justice to this component of the Environmental Genome Project, it is essential that sensitivity to these issues is upheld and that an effort is made to foster and maintain an open dialogue on these implications with both the scientific and nonscientific community.

One of the responsibilities faced by the Environmental Genome Project is to provide the scientific base upon which society can make better informed risk management decisions. We do not know nearly enough at the present time about how genetic susceptibility and environmental exposure collaborate in disease. The main goal of the Environmental Genome Project is to enhance population-based research toward identifying environmental exposure/disease relationships. As population susceptibilities are better understood, we will be in a better position than we are in today to make informed decisions about risk management. The approach proposed by the Environmental Genome Project offers great scientific opportunity and the potential to improve public health. To maximize our potential to enhance our health and our knowledge, we should remain open to new understanding and evolving technology or resources that might inspire a change in our approach to these important questions.

Samuel Wilson
NIHES
Research Triangle Park, North Carolina

Interregional Differences Undermine Sperm Trend Conclusions

The reanalysis of global trends in reported human sperm counts by Swan et al. (1) concluded that a decline in sperm densities was observed in the United States (1938–1988) and in Europe (1971–1990) but not in non-Western countries (1978–1989). The report notes that recent studies from Europe and the United States indicate large interregional differences in sperm density. Interregional differences noted in the United States (New York City vs. Los Angeles, CA) were as large as the reported differences in mean sperm density in 1938 versus 1990.

Regional heterogeneity should alert us to be cautious in interpreting temporal trends in reported sperm densities for each region (2). The only completely certain conclusion from the analysis of Swan et al. (1) is that there is a significant trend over time for sperm density studies to be reported from locations in the United States and Europe with lower sperm densities, while such a trend was not observed in reported studies from non-Western countries.

This limited conclusion is consistent with the data from single center studies where interregional differences are not a likely confounding factor. Single center studies in Europe report that sperm densities have declined over the last 10–20 years in Belgium (3), Finland (necropsy study) (4), London (area served by the Thames River water authority) (5), Paris (6), and Scotland (7) but not in Denmark (8), Finland (sperm count study) (9), London (outside area served by the Thames River water authority) (5), and Toulouse, France (10). Single center studies in the United States have reported no decline in sperm counts over the last 10–20 plus years in Los Angeles (11), New York City (11,12), Roseville, Minnesota (11), Seattle, Washington (13), and Wisconsin (14).

Given the inherent limitations in analysis of retrospective studies, prospective studies of human sperm counts are needed to determine trends in semen quality and to identify possible causes where temporal trends are observed (15).

John Heinze
John Adams Associates
Washington, D.C.

REFERENCES AND NOTES

1. Swan SH, Elkin EP, Fenster L. Have sperm densities declined? A reanalysis of global trend data. Environ Health Perspect 105:1228–1232 (1997).
2. Fisch H, Goluboff ET. Geographic variations in sperm counts: a potential cause of bias in studies of semen quality. Fertil Steril 65:1044–1046 (1996).
3. Van Waeleghem K, De Clercq N, Vermeulen L, Schoonjans F, Comhaire F. Deterioration of sperm quality in young healthy Belgian men. Human Reprod 11:325–329 (1996).
4. Loffredo, Silbergeld, and Parascandola.