CORRECTED 22 JUNE 2007; SEE LAST PAGE

MEDICINE

Environmental Biology and Human Disease

Better environmental biosensors are needed to study gene-environment interactions associated with disease.

David Schwartz* and Francis Collins

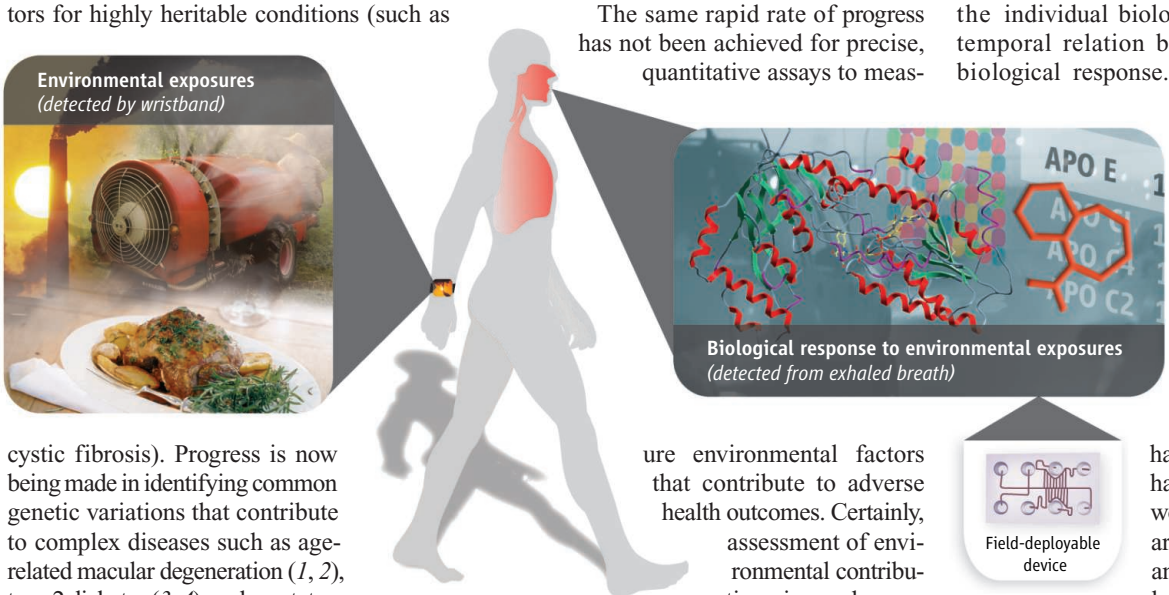
The etiology of most chronic human diseases (such as asthma, atherosclerosis, and cancer) is complex, involving a mix of genetic and environmental factors interacting with each other over hours, days, months, or years. Until recently, however, the disciplines of environmental sciences and genetics have proceeded independently; investigators in the former discipline have focused primarily on adverse conditions and diseases that are etiologically driven by environmental factors (such as benzene-induced leukemia), and those in the latter field have been finding genetic factors for highly heritable conditions (such as cystic fibrosis). Progress is now being made in identifying common genetic variations that contribute to complex diseases such as age-related macular degeneration (*1*, *2*), type 2 diabetes (*3*, *4*), and prostate cancer (*5*). However, the best opportunity to reduce risk in genetically susceptible people for the foreseeable future will not be to re-engineer their genes, but to modify their environment. The successful dietary treatment of phenylketonuria is a clear example.

We need to understand how genetic factors and environmental exposures interact in individuals to alter normal biological function and to affect the risk of disease development. This basic information is critical to understanding why and under what circumstances certain individuals develop disease and others remain healthy. Defining environmental contributions is also critical in identifying how and under what circumstances DNA sequence variations affect disease pathogenesis. For example, polymorphisms in CD14 and Toll-like receptor 4 are relevant in endotoxin-induced asthma but not in other more common types of asthma (*6*, *7*).

Progress in identifying genetic variations that contribute to common disease has been rapid in the last few years. Building on the foundation provided by the Human Genome Project, the International HapMap Consortium provided a public map (*8*) of human genetic variation. Dramatic advances in genotyping technology have led to a drop in cost of more than two orders of magnitude in just 5 years.

The same rapid rate of progress has not been achieved for precise, quantitative assays to measure environmental factors that contribute to adverse health outcomes. Certainly, assessment of environmental contributions is much more difficult than for genetic ones. The genome of an individual represents a bounded set of information, remains basically stable over time, and is very well suited to multiple analytical approaches. The potential universe of medically significant environmental exposures is much less well defined, and disease may appear several years after the exposure has ended. However, another explanation is apparent by contrasting the extensive investments in new genetic and genomic technologies over the past two decades with the much more modest expenditures in exposure sciences.

Traditional methods of assessing human exposure to chemical, dietary, physical, and psychosocial factors involve measuring the potentially toxic agent or exposure in environmental samples (air, water, or food) or biological specimens of blood and urine or, more commonly, characterizing the exposure event itself with regard to frequency, duration, and severity through questionnaires and other methods of recall. For example, the Centers for Disease Control and Prevention (*9*) have developed an extensive array of precise assays for toxins (natural source) and toxicants (synthetic source) that can be measured in various human specimens. However, these assays are not intended to provide information on the extent of the environmental exposure, the individual biological response, or the temporal relation between exposure and biological response. Existing methods of exposure assessment fail to capture the individual and dynamic extent of the exposure and its impact on fundamental biological processes.



Imagine that you could visit your family physician and be informed, by way of a personal sensor, that you have been exposed to a harmful substance that would explain why you are sick and inform you and your physician about how you should be treated. Imagine that you could wear a specialized wristband or a "smart shirt" that could alert you to the fact that your environment contains levels of air pollution that may increase your risk of having an asthma attack. Imagine that you had a hand-held device that could be used to determine whether the food you are eating contains harmful levels of trans fats, given your genetic predisposition to heart disease.

Recent advances in environmental and biological sensors suggest that the technologies are at hand, or can be readily engineered, to provide precise measures of chemical and biological hazards at the point of contact or to characterize the "biological fingerprint" left by a class of environmental stressors. Sensor technologies hold exceptional promise for providing critical information for continuous (real-time) data

CREDIT: ADAPTED BY N. KEVITIYAGALA/SCIENCE; (PHOTOS) GETTY IMAGES; (PROTEIN STRUCTURE) NIST

collection and simultaneous measurement of multiple agents (multiplexing) within a single device. New sensing modalities have emerged from nanotechnology and nanoengineering, medical diagnostics, and biodefense arenas that could be adapted and developed for the exposure sciences. Artificial receptors such as molecularly imprinted polymers (MIPs) have potential as stable surrogates for biological recognition agents such as antibodies, enzymes, tissues, or cells (*10*).

In addition to being self-contained, these sensors should be capable of quantitative, continuous data capture in the field, without the need for sample processing and analysis at a laboratory. These devices must be easy-to-use; portable; minimally inconvenient (wristband, watch, phone, or lightweight purse); rugged; and inexpensive to deploy. This will require sophisticated computer systems and analytic approaches that can handle the immense volume and complexity of data generated for each individual and, also, would allow for integration of data on environmental exposures with genetic factors for the individual and the population.

Multiple molecular changes can result from environmental stressors, but not all of these changes are linked to increased disease risk. Thus, biosensors will need to record internal, molecular events that signify increased risk of disease from exposures to different forms of environmental stress, such as patterns of gene or protein expression (*11*), as well as to measure response indicators such as DNA or protein adducts (*12*) that persist even after the exposure has ended.

The importance of this opportunity has been recently highlighted at the National Institutes of Health by the launching of the Genes, Environment, and Health Initiative (GEI) (*13*), with strong support from U.S. Health and Human Services Secretary Michael Leavitt and NIH Director Elias Zerhouni. This $40 million-a-year interdisciplinary initiative, managed by a coordinating committee that we cochair, includes an Exposure Biology Program. The near-term goal of the program is to develop new noninvasive tools and biomarkers for assessing individual exposures to environmental stressors that interact with genetic variation to result in human disease. However, to fully appreciate the predictive importance of these measures of exposure, this technology needs to be deployed in large-scale case-control and population-based genetic studies of health and disease, some of which will include genome-wide association analysis through support from the Genetics Program of GEI and the National Children's Study (*14*).

Establishing partnerships in the scientific, technical, engineering, and business communities will be critical to our success. We need to define collectively appropriate milestones and deliverables for what can be achieved with this new technology, focusing on the types of environmental exposures and response indicators to be measured, the type of applications, and the level of temporal and spatial resolution. For example, there is a vast range of potential applications for small-scale sensing devices, such as cell-based microsystems or lab-on-a-chip technology, from the detection of individual molecules within single cells, to the measurement of global changes in genes, proteins, and metabolites in peripheral biofluids (*15*, *16*). We need to define the most critical questions in exposure biology up front so that scientists and the public understand the scope of technology that is required. In turn, engineers and manufacturers must identify the technological limitations and needs in this field.

Short-term strategies should target specific, attainable goals and deliverable devices that provide integrated panels of biomarkers for priority classes of environmental stressors, such as pesticides and solvents, as well as cholesterol-rich or heavy metal–contaminated foods. Immediately available technologies include point-of-contact environmental sensors and biosensors based on molecular assays, such as global protein or metabolite profiling and molecular imaging. Short-term strategies could be adapted for future applications that provide a much broader range of analytes and that include measurement of previously unknown stressors. More long-term investments should target high-risk, potentially high-benefit technologies, such as lab-on-a-chip or microfluidic devices, molecular probes, and imaging systems that incorporate multiplexed sensing capabilities for concurrent detection and quantification of environmental stressors with geospatial referencing and remote, real-time data capture.

A major hurdle is that the expertise needed to advance the science spans so many highly specialized fields and that interdisciplinary training and research opportunities are just beginning to evolve. In that regard, the National Institute for Environmental Health Sciences and the National Human Genome Research Institute have initiated a training program in environmental genomics (*17*) to bring these disciplines closer together and to train a new generation of scientists who are equally at home in both fields.

Initial field deployment will need to focus on small-scale studies and leverage, to the extent possible, existing population studies of environmental and genetic risk factors. This will allow time and opportunity to develop protocols that standardize methods of sample collection, processing, and storage; labeling molecules such as peptides with isotope tags; developing internal reference standards; and conducting data analysis. These steps are essential to achieve sufficient reproducibility and reliability of results that would make large-scale studies worthwhile. Concurrently, there would also be tremendous value to developing a biosample and data repository to promote sharing of scientific resources and discoveries made in exposure biology across multiple research programs. Of course, this would require a bioinformatics infrastructure to analyze the volumes of environmental and biomarker data generated and to integrate these data with the corresponding genomic information available for each individual and/or population. Ultimately, it is hoped that GEI will provide a way to identify subsets of individuals with high disease risks due to particular combinations of genetic variants and environmental exposures or stressors, as well as to lead to targeted therapies and intervention techniques for disease prevention and more effective health maintenance.

References and Notes

1. R. J. Klein *et al.*, *Science* **308**, 385 (2005).
2. Z. Yang *et al.*, *Science* **314**, 992 (2006).
3. S. F. Grant *et al.*, *Nat. Genet.* **38**, 320 (2006).
4. R. Sladek *et al.*, Nature **445**, 881 (2007).
5. L. T. Amundadottir *et al.*, *Nat. Genet.* **38**, 652 (2006).
6. F. D. Martinez, *Immunol. Allergy Clin. North Am.* **25**, 709 (2005).
7. N. C. Arbour *et al.*, *Nat. Genet.* **25**, 187 (2000).
8. International HapMap Consortium, *Nature* **437**, 1299 (2005).
9. National Center for Environmental Health, www.cdc.gov/nceh/.
10. V. B. Kandimalla, H. Ju, *Anal. Bioanal. Chem.* **380**, 587 (2004).
11. A. Spira *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10143 (2004).
12. J. D. Groopman *et al.*, *Cancer Res.* **52**, 45 (1992).
13. Genes, Environment, and Health Initiative, http://genesandenvironment.nih.gov/.
14. National Children's Study, http://nationalchildrensstudy.gov.
15. J. El-Ali, P. K. Sorger, K. F. Jensen, *Nature* **442**, 403 (2006).
16. H. Craighead, *Nature* **442**, 387 (2006).
17. Training program in environmental genomics, http://staging.gei.nih.gov/traininggrants.asp.

18. Many of the ideas discussed in this manuscript evolved from the NIH Exposure Biology Workshop sponsored by National Institute of Environmental Health Sciences and National Human Genome Research Institute held in Greensboro, North Carolina, in May 2006 (www.gei.nih.gov/exposurebiology/meetings/exposurebiology06/). We thank the workshop participants for their commitment to this evolving field and B. Weis for her thoughtful advice and critical contributions to this manuscript.

10.1126/science.1141331

ERRATUM

Post date 22 June 2007

Policy Forum: "Environmental biology and human disease" by D. Schwartz and F. Collins (4 May 2007, p. 695). The authors' affiliations and contact information were omitted. The authors are with the National Institutes of Health, Bethesda, MD 20892, USA. David Schwartz is the author for correspondence. E-mail: david.schwartz@niehs.nih.gov