

HUMAN GENETICS

Environment Institute Lays Plans for Gene Hunt

BETHESDA, MARYLAND—Twelve people died after members of the Aum Shinrikyo cult unleashed a potent nerve gas called sarin in the Tokyo subway 2 years ago. Some of these victims, scientists now know, may have been much more vulnerable than others. Circulating in the blood of 25% of Asians and 10% of Caucasians is a version of the enzyme paraoxonase that converts sarin to a less toxic chemical about 10 times more quickly than the enzyme found in most people.

The paraoxonase gene is one of dozens that toxicologists think make some individuals more susceptible to the effects of pollutants and other environmental chemicals, contributing to everything from cancer to birth defects and Parkinson's disease. Hoping to ferret out dozens more of these "environmental susceptibility" gene variants, National Institutes of Health (NIH) scientists are putting together a major effort to sequence DNA from perhaps 1000 people to try to demonstrate a link between certain genes and patterns of disease. "This is information that can really revolutionize public health policy" by making it possible to identify and protect people susceptible to hazards, says Ken Olden, director of the National Institute of Environmental Health Sciences (NIEHS), whose scientists conceived the so-called Environmental Genome Project.

But the path to this revolution is far from certain. Top genome experts and population geneticists invited to an NIEHS symposium here last week to discuss the project's feasibility pointed to many potential obstacles. First, there's the novelty factor: No genome project has yet attempted to survey genetic diversity for a large number of human disease genes. Then there's the expertise factor: NIEHS, an NIH branch in Research Triangle Park, North Carolina, is best known for toxicology—not genomics—and will need plenty of help from the rest of the scientific community, some experts cautioned. And finally, there's the turf factor: National Human Genome Research Institute director Francis

Collins has outlined an ambitious proposal to create a public data bank of so-called single nucleotide polymorphisms, or SNPs—minute variations in sequenced genes (*Science*, 19 September, p. 1752)—that might overlap with NIEHS's project. "There's a lot of unknowns in this endeavor, and we have to proceed cautiously," acknowledges NIEHS scientific director Carl Barrett. But the project's uncertainties shouldn't hold it back, says symposium co-organizer Lee Hartwell, head of the Fred Hutchinson Cancer Research Center in Seattle. "There's no reason why we can't get started."

The idea for the undertaking follows several decades of work on common variations of genes involved in activating or detoxifying drugs and chemicals that we breathe, drink, or eat. "Each person basically has his own unique fingerprint of drug-metabolizing enzymes and receptors, so we all handle drugs [and chemicals] differently," says Dan Nebert, director of the Center for Ecogenetics and Environmen-

NAT variants may have up to a sixfold greater risk of bladder cancer than nonsmokers with other NAT variants. (These cancer susceptibility genes will be discussed in a special report on cancer in the 7 November issue of *Science*.)

Such variants appear to be much more common and less dangerous than mutations in genes such as the breast cancer gene *BRCA1* that increase disease risk dramatically, seemingly independently of environmental stimuli. But the sheer prevalence of *P450* and other variants could result in a major population risk. "If the allele is common," explains NIEHS molecular epidemiologist Jack Taylor, "it can account for an incredibly large fraction of disease in the population."

So far, researchers have managed to indict only a handful of environmental risk genes. In just a few cases are there "really solid human studies that truly demonstrate" a link between one of these genes and disease, says toxicologist John Groopman of Johns Hopkins University. Carefully sifting through the genome to find new candidates, Taylor and others say, could have profound implications for encouraging susceptible people to avoid certain exposures and by setting safer standards for workers and the public (see sidebar).

But NIEHS officials are unsure about how to begin the hunt and which genes to target first. At the meeting, Barrett sketched out

NIEHS's preliminary thinking: First, scientists would set up a DNA repository from 1000 individuals representative of the major U.S. ethnic groups. Next, teams would use some combination of DNA chips and sequencers to resequence alleles of roughly 200 "candidate" genes involved in everything from DNA repair to digestion (see table). At about 10 cents per nucleotide, NIEHS says this step could cost \$200 million. Next, researchers would sort out which alleles are common enough to be classified as polymorphisms—versions shared by 1% or more of the population. The project would then sponsor molecular, animal, and, finally, population studies—of sick

people who had been exposed to a suspect chemical, for example—to find out how important these polymorphisms are to disease.

Although experts at the meeting endorsed the project's overall goals, they found plenty to fault in the game plan itself. For instance, some meeting participants said, NIEHS may have greatly underestimated the project's price tag, which relies on savings



A SAMPLING OF ENVIRONMENTAL GENES

Polymorphism	Class	Environmental exposure	Associated disease
CYP1A1	Activation	Smoking	Lung cancer
NAT2	Detoxification	Smoking	Bladder, breast cancer
GSTT1 (null)	Detoxification	Chlorinated solvents	Cancer, toxicity
Paraoxonase	Detoxification	Nerve agents, pesticides	Nervous system damage
HLA-H	Nutritional factors	Iron in diet	Hemochromatosis
TGF- α	Growth factor	Maternal smoking	Cleft lip & palate
Locus on chrom. 17 in mice	Immune/inflammatory response	Ozone	Lung inflammation
HLA-DP bet1 marker	Immune response	Beryllium	Chronic beryllium disease (lung disorder)
ALAD	Biosynthesis	Lead	Lead poisoning

tal Science at the University of Cincinnati.

Variations, or alleles, of the gene for paraoxonase—an enzyme that breaks down toxic organophosphate compounds, including many insecticides—is one example. Many other genes, such as those in the cytochrome P450 and NAT families—which metabolize carcinogens—can increase cancer risk, especially in smokers. For instance, a smoker with certain

A More Rational Approach to Gauging Environmental Threats?

One of the biggest potential payoffs from an environmental genome project (see main text) is that it could help policy-makers devise rules that better protect sensitive individuals. "To have intelligent environmental regulatory policy, one has to begin to unravel the role of genetics in determining the differences in susceptibility," says National Institute of Environmental Health Sciences (NIEHS) director Ken Olden.

Olden's words are music to the ears of members of Congress who have been clamoring for better science behind regulations. Risk assessors at the Environmental Protection Agency (EPA) and elsewhere now craft rules with a standard fudge factor to try to protect sensitive individuals: They set the permissible exposure level to a chemical, for instance, at a tenth of that deemed acceptable for the general population. Data on the prevalence of susceptibility genes could reduce the need for guesswork, says NIEHS's George Lucier, who's helping write the EPA's dioxin reassessment. "As we get more and more information on the variation of environmentally relevant genes across the population," he says, "we'll be able to more frequently ... use real data."

In some cases real data might result in a less stringent standard and in others a tighter one. For example, some people may have a version of a detoxifying enzyme that makes them five times more sensitive than others to a pollutant. Risk assessors, then, might permit an exposure that's a fifth that of the acceptable level for the rest of the population—an exposure twice as high as they might set using the standard fudge factor. "It could go either direction," Lucier says. "It could be a 100-fold factor or a twofold factor." But political factors and economics, in some cases, will inevitably pull rank on science—particularly if the sensitive population is tiny. "It would obviously become extremely expensive to protect a few individuals," says University of Washington, Seattle, toxicologist Dave Eaton.



New direction. The time is right for NIEHS to search for susceptibility genes, says Olden.

DUANE HALL

Indeed, both scientists and regulators may struggle to "digest and understand the meaning and importance" of the initial data on environmental genes, says George Gray of the Harvard Center for Risk Analysis. For one thing, several genes may be involved in defining an individual's risk. Take, for instance, benzene, which at high exposures has been linked to human leukemia. After sampling liver tissue from 10 people, a team led by toxicologist Michele Medinsky of the Chemical Industry Institute of Technology in Research Triangle Park, North Carolina, found a 14-fold difference across the samples in the activity of CYP2E1, an enzyme that converts benzene into chromosome-damaging metabolites. If this variation is due to genetics, says Medinsky, then people with lower CYP2E1 activity would be "relatively protected" from benzene toxicity. But because people also vary in their activity levels of other enzymes that detoxify the benzene metabolites, the risk for someone with fast-acting CYP2E1, she asserts, "in fact might not be elevated at all."

Another bedeviling issue is how this information might be used to alter workplace exposure levels. One test case might be beryllium, an industrial metal that can cause an incurable lung disease. Four years ago, scientists found a genetic marker of susceptibility to beryllium disease carried by 30% of the population; 97% of a group of workers with the disease had the marker. Employers are now debating whether to screen workers for it. One worry is that a worker with a susceptibility gene could be denied a job.

The trend in risk assessment is to factor in genetic susceptibility by developing "ranges and distributions" of risk, rather than a single number—leaving it to managers to work out how to use the information, says Gray. EPA's proposed cancer risk guidelines encourage this kind of analysis, but don't specify how to do it. Says Lucier, "I think the regulatory agencies need to really start getting their thoughts together about how this information will be used." —J.K.

from untested technologies such as the DNA chips for resequencing being developed by the biotech firm Affymetrix Inc. of Santa Clara, California. Also unclear is how much to sequence—whether regulatory regions should be included, for example.

The biggest hurdle, however, is the crucial step that follows sequencing: determining the relevance of polymorphisms to disease. The key questions, says Barrett, are "how much variation to expect and whether you can distinguish important changes from unimportant ones." Scientists argue that any effort to find suspect polymorphisms is complicated by the often dizzying variation of alleles. "We should not be too facile in thinking we can find causal sites," says Penn State geneticist Ken Weiss. Adds National Cancer Institute clinical sciences director Ed Liu: "I'm increasingly impressed with the complexity of what's involved and how little data we have that would guide us."

The bottom line, experts say, is that

NIEHS shouldn't try to go it alone—particularly in the early stages of sample gathering. "If we end up with several different efforts for collecting samples, that will really be too bad," says Collins, who urged NIEHS to collaborate on the SNPs database. Indeed, Barrett says, NIEHS is now likely to join that effort, which, as Collins outlined at the meeting, would aim to assemble a DNA repository representative of at least some of the U.S. population. NIH will host a meeting on 8 and 9 December to decide "how to design the sample set," Collins says. In the meantime, NIEHS plans to start compiling a list of candidate genes. (NIEHS is soliciting suggestions at: www.niehs.nih.gov/diroscd/policy/egp/home.htm)

Even more uncertain is how NIEHS will move from genotype to phenotype. "Everybody realizes we really don't know how to do it yet," Hartwell says. One way to approach this, suggests Liu, would be to start small—by resequencing, say, just 10 genes—and assess how well functional and epidemiological studies

succeed at tying variations in those genes to disease before proceeding to the next batch of genes. Another idea is to pick a gene whose variance is already well understood—NAT2, for example—and characterize its alleles from scratch using DNA repository samples to see how approaching the problem from the genomic end compares to the earlier toxicology and molecular epidemiology studies.

According to Collins, "it's going to be a few months to see how this shakes out." Barrett agrees and says that NIEHS may delay putting out its initial request for grant proposals for \$10 million worth of seed money until late next year, after key details of the project's design are worked out. There's no doubt, however, that NIEHS's grand plans have got others in the community revved up about environmental genes. "NIEHS is way out in front," says genome sequencer Glen Evans of the University of Texas Southwestern Medical Center in Dallas. "This is kind of the wave of the future."

—Jocelyn Kaiser