cataway, New Jersey, plans to pull out about 100 genes using a transposon called *Activator*. At Cold Spring Harbor Laboratory in New York, Rob Martienssen will use a transposon system to create some 40,000 lines of corn mutants, each associated with a piece of altered DNA. Although his approach is similar to Walbot's, "having several groups use complementary approaches increases the likelihood that every important gene will be identified," says NSF's David Meinke.

A view across the field

Corn is the only crop plant in which transposons can be easily used to pull out genes. But researchers think they will be able to combine what they learn about the corn genome with data coming from the *Arabidopsis* sequencing project and also a rice genome project expected to be under way in Japan and other countries within the year (see sidebar on previous page). There appear to be enough similarities between the gene arrangements in different species that locating a particular gene in one will point to counterparts in the others. But first, says Cornell University plant molecular geneticist Steven Tanksley, "we need to find ways to connect [*Arabidopsis* and rice] genome information to other species."

To find those connections, Andrew Paterson, a plant molecular geneticist at Texas A&M University in College Station who is moving in January to the University of Georgia, Athens, will look for similar DNA landmarks in sorghum, rice, and corn. And Tanksley's team will be looking at genes involved in fruit development in wild and domestic tomato plants and comparing them with *Arabidopsis*, with an eye to evaluating how evolution has reshaped genomes. "All of these factors will merge into a picture of the interrelatedness that will tie one crop to another," Coe says.

While these groups are exploring the fundamental structure of plant genomes, others will jump into functional genomics—determining how patterns of gene expression vary under different conditions or in different mutants. Among other things, this should help identify genes that affect plant yields or responses to stresses such as high salt concentrations or infection by pathogens.

For example, plant geneticist Bertrand Lemieux of the University of Delaware, Newark, wants to find the genes that enable some corn varieties to produce more oil than others, and UC Davis's Wilkins will try to track down all the genes important to cotton fiber formation—information that could ultimately lead to improved cotton varieties. A team coordinated by Hans Bohnert, a biochemist at the University of Arizona, Tucson, will focus on identifying genes involved in salt tolerance, while Nina Fedoroff of Pennsylvania State University in University Park and her colleagues will look for genes that turn on or increase their activity when plants are subjected to high concentrations of ozone and damage by pathogens. "Rather than providing just [DNA] sequence, we're attacking a biological problem," Bohnert says.

Once identified, such genes might be used to genetically engineer plants with improved yields or resistance to the various stresses. Fedoroff hopes eventually to create inexpensive monitors that will let farmers detect when their crops are at risk. It may take years to achieve these goals, Fedoroff and Bohnert note. But in the meantime these projects will invigorate basic research. Genes involved in fiber formation, for example, will help plant physiologists understand cell growth in general, and there should soon be a wealth of new genes of all kinds to study in corn. Says Tanksley, "plant biology, like all biology, has embarked on a golden age." Or, as Gerald Tumbleson, a Minnesota corn farmer, said at a press conference announcing the NSF awards, "With this season of biology, we're going to be able to do things that we only dreamed of before. I just wish I was 20 years old, because I think this is fantastic." –ELIZABETH PENNISI

DNA Studies Challenge the Meaning of Race

NEWS

Genetic diversity appears to be a continuum, with no clear breaks delineating racial groups

Last year, the U.S. Office of Management and Budget (OMB) completed a contentious 4-year review of the racial and ethnic categories that will be used to define the U.S. population in federal reports, including the 2000 census. It finally settled on seven groupings: American Indian or Alaska Native; Asian; Black or African American; Native Hawaiian (added after OMB received 7000 postcards from Hawaiians) or Other Pacific Islander; White; Hispanic or Latino; and Not Hispanic or Latino. The categories could have enormous implications—from the distribution of government resources to political districting to demographic research. But as far as geneticists are concerned, they're meaningless.

"Ridiculous" is the word cultural anthropologist John Moore of the University of Florida, Gainesville, uses to describe such racial typing. This view is based on a growing body of data that indicates, as Moore says, that "there aren't any boundaries between races." Geneticist Kenneth Kidd of Yale University says the DNA samples he's examined show that there is "a virtual continuum of genetic variation" around the world. "There's no place where you can draw a line and say there's a major difference on one side of the line from what's on the other side." If one is talking about a distinct, discrete, identifiable population, Kidd adds, "there's no such thing as race in [modern] *Homo sapiens*." Indeed, the American Anthropological Association urged the government last year to do away with racial categories and, in political matters, let people define their own ethnicity.

You might think that this emerging view of genetic variation would help lower the temperature of discussions about race and ethnicity. But, ironically, researchers who want to extend their studies of genetic diversity are being stymied by the intense sensitivity surrounding the topic. A major international project to survey genetic diversity around the globe is on hold, having been opposed by activists. Moreover, a planned database of genetic polymorphisms is being constructed in a way that will prevent comparisons between different population groups, making it useless for exploring the gene frequency variations that do exist, according to researchers.

Anthropologists have long objected to the stereotypes that are used to classify human populations into racial groups. But the most potent challenge to such groupings has come from genetic studies of human origins. The field was "transformed" in the late 1980s, says anthropologist Kenneth Weiss of Pennsylvania State University in University Park, by an analysis of variations in mitochondrial DNA (mtDNA) begun by Rebecca Cann of the University of Hawaii, Manoa, Mark Stoneking of Penn State, and the late Allan Wilson of the University of California, Berkeley. These researchers reported that diversity in mtDNA genes was two to three times greater in Africa than in Europe or the rest of the world. Assuming that the rate of change in mtDNA was fairly constant, they concluded that Africans' mtDNA was older than that of non-Africans, and that modern humans originated from a small population that emerged from Africa and migrated around the globe.

Since the 1980s, other researchers have extended these studies by looking at diversity in nuclear DNA. Two years ago, for example, Kidd and his Yale colleague Sarah Tishkoff reported patterns of variation in the CD4 gene locus on chromosome 12 among 1600 individuals chosen from 42 populations from around the world (*Science*, 8 March 1996, p. 1380). They have since looked at 45 short tandem repeats across the entire nuclear genome in multiple populations. What they found, says Kidd, is "a lot of genetic variation in Africa, decreasing genetic variation as you go from west to east across Eura-

sia, and decreasing more into the Pacific, and separately decreasing into North America and South America." The best explanation for this pattern, Kidd argues, is the same one Wilson and his colleagues put forward: A small group of people moved out of northern Africa to colonize the rest of the world.

Not all the studies of nuclear DNA have been consistent. For example, research on the Y chromosome by Michael Hammer of the University of Arizona, Tucson, and others found greater genetic diversity for this chromosome in Asia than in Africa. And some studies—such as those by Henry Harpending and colleagues at the University of Utah, Salt Lake City—have found that the difference between Africa and the rest of the world in the amount of variation in nuclear DNA is much smaller than reported for mtDNA. But the new data generally point in the same direction: Human genetic diversity is greatest in Africa, and the genetic heritage of modern humans is largely African.

Although researchers are far from unanimous, even some who have been cautious about interpreting regional patterns of mtDNA variation seem ready to accept the out-of-Africa thesis today. Lynn Jorde, a colleague of Harpending at the University of Utah, says, "All of us have been a little suspicious of the mitochondrial DNA data because it is such a small part of our genome." But the totality of evidence—particularly studies showing that common variants found outside Africa are mainly "subsets" of those in Africa—persuade Jorde that the out-of-Africa theory is right. Indeed, Jorde says, this hypothesis was accepted by most of the 100 or so geneticists, physical anthropologists, and linguists attending a conference on human origins at Cambridge University, in Cambridge, U.K., last month. "There's enough agreement to tell us we're on the right track, but enough disagreement to keep things interesting," Jorde says.

CREDIT: JEFF GREENBERG/VISUALS UNLIMITED

In 1991, population geneticist L. Luca Cavalli-Sforza of Stanford University and his colleagues proposed an ambitious plan to probe this genetic continuity more deeply by collecting and analyzing DNA samples from thousands of populations around the globe. The effort, called the Human Genome Diversity Project (HGDP), ran into heavy fire on ethical grounds and fears that it might violate indigenous peoples' rights, however, and it remains stalled for lack of funds (*Science*, 24 October 1997, p. 568).

With the HGDP mired in controversy, many population geneticists were hoping that some of its goals might be achieved through a database proposed by the National Human Genome Research Institute (NHGRI) of the National Institutes of Health. In January, NHGRI director Francis Collins launched a scheme to create a collection of samples of human DNA containing point mutations called single nucleotide polymorphisms (SNPs), which can be used as markers in the study of inherited diseases. This $30 million project aims to gather a set of 100,000 SNPs representing human diversity (*Science*, 19 December 1997, p. 2046).

The plan ran into difficulties when officials had to decide which populations should be included in the 450 DNA samples to be analyzed for SNPs. Federal guidelines then recognized four races (black, white, Asian or Pacific islander, and Native American or Alaskan native), as well as an ethnic type (Hispanic). But because the categories aren't based on genetics, NHGRI sought advice on how to structure its groupings from Weiss and anthropologist Jonathan Friedlaender at Temple University in Philadelphia, among others.

Weiss and Friedlaender say the federal race and ethnic categories are useless for a scientific sampling program. "Take a term like 'Hispanic,' " says Weiss: "It's biologically a very bad term," because it lumps together people from Cuba, Puerto Rico, and Mexico who have fundamentally different histories. "So you're labeling people the wrong way" if you try to use these labels for a balanced DNA collection, says Weiss.

In the end, NHGRI opted for a quick solution. It is using samples from U.S. residents grouped under broad geographical ancestry headings: African, Asian, European, North and South American. For convenience, nearly all the samples are coming from an existing set of DNAs, supplemented with some extra samples from Native Americans. "This is a first step," says program director Lisa Brooks: "It covers a huge amount of human variation without claiming to cover everything."

> **"One of the benefits that's going to come from [studies of genome diversity] is an even greater understanding of how similar we all are in our marvelous variation."**
>
> **—Kenneth Kidd**

Those categories might provide a basis for population comparisons, but NHGRI made a second broad decision that, according to some scientists, will preclude such studies. "We're not identifying who these individuals are [in the SNP database] by ethnicity, or sex, or anything else," says Brooks. She adds that, "We've gone to great pains to ensure that people who use these resources will not identify ethnicity" of the DNA they study. Research on alcoholism or schizophrenia, for example, could cause offense if linked to a specific group, and NHGRI wants to avoid any "group stigmatization."

"As far as I'm concerned," the removal of population source data from a DNA sample "means the sample is useless," says Kidd. "I won't use it." Kidd insists that genetic markers such as SNPs are valuable only if they can be understood within the context of the population from which they're drawn, and for this, one must know the source. Florida's Moore agrees. Making the SNPs completely anonymous "drives a wedge" between anthropology and the new genetic database, he says. Kidd says he will have to rely on his own 15-year-old collection, which includes the DNA source information. Cavalli-Sforza's group and others are also making do with independent data collections to continue their own, small-scale versions of the HGDP, tracing broad patterns of human genetic variation.

At a meeting this year, Kidd predicted, "One of the benefits that's going to come from [studies of genome diversity] is an even greater understanding of how similar we all are in our marvelous variation." For now, however, that's still a dream. **–ELIOT MARSHALL**