

Back to the
Future
of
Cereals

Genomic studies of the world's major grain crops, together with a technology called marker-assisted breeding, could yield a new green revolution

BY STEPHEN A. GOFF
AND JOHN M. SALMERON

RICE SEEDLINGS can be genetically tested for desirable traits.



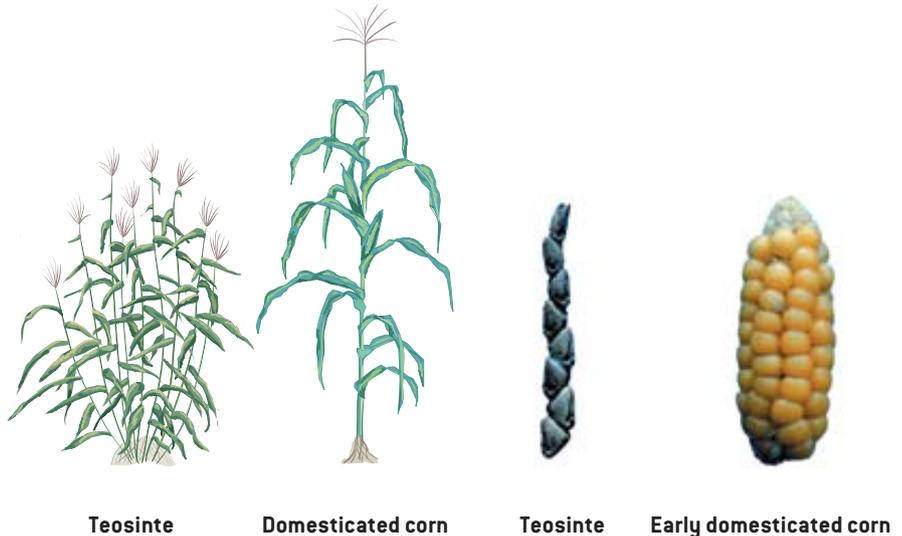
For thousands of years,

farmers have surveyed their fields and eyed the sky, hoping for good weather and a bumper crop. And when they found particular plants that fared well even in bad weather, were especially prolific, or resisted disease that destroyed neighboring crops, they naturally tried to capture those desirable traits by crossbreeding them into other plants. But it has always been a game of hit or miss. Unable to look inside the plants and know exactly what was producing their favorable characteristics, one could only mix and match plants and hope for the best.

Despite the method's inherent randomness, it has worked remarkably well. When our hunter-gatherer ancestors started settling down some 10,000 years ago, their development of agriculture allowed human society to undergo a population explosion. It is still expanding, demanding continual increases in agricultural productivity.

Yet 99 percent of today's agricultural production depends on only 24 different domesticated plant species. Of those, rice, wheat and corn account for most of the world's caloric intake. Each of these three extremely important cereals is already produced in amounts exceeding half a billion tons every year. To keep pace with a global population projected to reach nine billion by 2050, while maintaining our present average daily consumption of between 0.9 and 3.3 pounds of these grains per person, cereal crops will have to yield 1.5 percent more food every year and on a diminishing supply of cultivated land.

Plant scientists believe that crop yields have not yet reached their theoretical maximum, but finding ways to achieve that potential increase and to push the



MODERN CORN AND ITS ANCESTOR TEOSINTE look so dissimilar [drawings] that their relationship was questioned until genetic investigations confirmed it. By selectively propagating plants with desirable traits, ancient cultivators in what is now Mexico unwittingly favored certain versions of genes that control branching pattern, kernel structure and other attributes. By 4,400 years ago the teosinte cob's hard fruit case (left photograph) was gone and plump, modern-looking corn cobs (right photograph) carried the versions of the genes that control protein storage and starch quality in all domesticated corn today.

yield frontier still further is an ongoing international effort. Encouragingly, a new set of tools is revealing that some of the answers may be found by exploring the origins of the three major cereal crops.

Creating Modern Crops

MOLECULAR AND GENETIC studies are showing that wheat, rice and corn, as well as barley, millet, sorghum and other grasses, are far more interrelated than was once thought, so fresh insights into any one of these crop species can help improve the others. Further, many of these improvements may come from tapping the genetic wealth of our crops' wild ancestors by breeding useful traits back into the modern varieties.

Although the cereal crops are descendants of a common ancestral grass, they diverged from one another some 50 million to 70 million years ago, coming to inhabit geographically distinct regions of

the world. Beginning around 10,000 years ago, farmers in the Mediterranean's Fertile Crescent are believed to have first domesticated wheat, and perhaps 1,000 years later, in what is now Mexico, farmers began cultivating an ancestor of modern-day corn. The ancient Chinese domesticated rice more than 8,000 years ago.

As our ancestors domesticated these plants, they were creating the crops we know now through a process very much like modern plant breeding. From the wild varieties, they selectively propagated and crossbred individual plants possessing desirable traits, such as bigger grains or larger numbers of grains. Plants that did not disperse their seeds were appealing, because harvesting their grain was easier, although this characteristic made a plant's propagation dependent on humans. Early cultivators also selected plants for their nutritional qualities, such as seeds with thin coats that could be eaten easily and maize varieties whose starch consistency best lent itself to making tortillas. In this way, crop plants became increasingly distinct from their progenitors and eventually rarely crossed with their wild versions. Corn became so dissimilar to its ancestor, teosinte, that its origin was commonly disputed until very recently [see illustration above].

This human modification of cereal plants through selective propagation and

Overview/Tapping Crops' Genetic Wealth

- Comparing the genomes of major cereal crop species shows their close interrelationships and reveals the hand of humans in directing their evolution.
- Identifying the functions of individual plant genes allows scientists to search modern crops and their wild relatives for gene versions that confer desirable traits.
- With the desired gene as a traceable marker, traditional crossbreeding can become faster and more precise.

crossbreeding begun during prehistoric times has never stopped. Over the past century, crops have been selected for larger seed-bearing heads to increase their yields. These higher-yielding seed heads are heavy, so shorter plant heights were also bred into rice and wheat to prevent the plants from being bent to the ground by wind. Breeding for disease resistance, environmental stress tolerance and more efficient utilization of nitrogen fertilizers dramatically increased yields and their consistency, producing the green revolution of the 1960s. Corn's average yield per acre in the U.S., for example, has risen by nearly 400 percent since 1950.

Yet even during that boom period, plant breeders had little more to go on than the earliest crop cultivators. Most were limited to visible plant characteristics, or markers, such as seed size or plant architecture, to guide their selection of desirable lines for further propagation.

Still, studies of the genomes of cereal crops illustrate how prehistoric cultivators, by selecting for visible traits, were unwittingly selecting particular genes. For example, a group led by Svante Pääbo of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, analyzed the alleles, or versions, of specific genes in corn cobs recovered from sites in Mexico near the origin of corn domestication. Pääbo and his colleagues determined that by 4,400 years ago, domesticated corn already possessed genetic alleles that control the plant's branching pattern as well as aspects of protein and starch quality found in all modern corn varieties. In corn's wild relative, teosinte, these alleles occur in only 7 to 36 percent of plants, indicating that the selection pressure applied by early farmers to favor those alleles was rapid and thorough.

Indeed, working independently on different cereal crop species, breeders have been unknowingly altering them by selecting mutations in similar sets of genes. Trait mapping—narrowing the probable location of the gene underlying a trait to a particular chromosomal region, or locus—has shown that many of the changes humans have made in modern cereals map to similar loci in the genomes of related crop plants. The reason for this similarity

is that the structures of these different crops' genomes are themselves so similar, despite millions of years of independent evolution separating the cereal species.

Harvesting Genomes

A FEW THOUSAND trait-controlling loci have now been mapped in various domesticated cereals, revealing the surprising degree to which the plants' overall genetic maps have been conserved. The high degree of this correspondence, known as synteny, between genomes of all the grasses allows scientists to consider them as a single genetic system, meaning that any discoveries of genes or their function in one cereal crop could help scientists to understand and improve the others.

Rice, whose formal name is *Oryza sativa*, is likely to be the first to yield many of these new insights, because it will be the first crop plant to have its entire genome sequenced. One of us (Goff) has already published a draft sequence of the *japonica* subspecies of rice most commonly grown in Japan and the U.S., and Chinese researchers have produced a draft of the *indica* subspecies widely cultivated in Asia. The International Rice Genome Sequencing Project is expected

to complete a detailed sequence of rice's 12 chromosomes by the end of this year.

The rice genome is the easiest of all the cereals' to tackle because it is much smaller than the others, with only 430 million pairs of DNA nucleotides. By comparison, the human genome has three billion of these so-called base pairs, as does corn. Barley's genome contains five billion base pairs and wheat, a whopping 16 billion. A corn genome-sequencing project is under way, and one for wheat is under consideration. And from the existing sequence information about rice, tens of thousands of genes have already been identified. Just knowing that a stretch of the genome is a gene does not tell us what it does, though.

Several strategies allow us to determine a gene's function, but the most straightforward involves searching existing databases of all known genes to look for a match. Often genes are responsible for such basic cellular activities that a nearly identical gene will be found in microbes or other organisms whose genes have already been studied. Of the 30,000 to 50,000 predicted genes in rice, approximately 20,000 have sequence similarity, or homology, to previously discovered genes whose function is known,

DESIRABLE TRAITS

Traits that plant breeders seek to modify fall into broad categories, including growth, plant architecture, stress tolerance and nutrient content. Yield increases—the holy grail of agriculture—can be achieved by expanding the size or number of grains produced by a single plant, by enabling more plants to grow in the space usually needed for one, or by making plants tolerant of conditions where they previously could not thrive.

Growth

Grain size or number
Seed-head size
Maturation speed

Architecture

Height
Branching
Flowering

Stress tolerance

Drought
Pests
Disease
Herbicides
Intensive fertilization

Nutrient content/quality

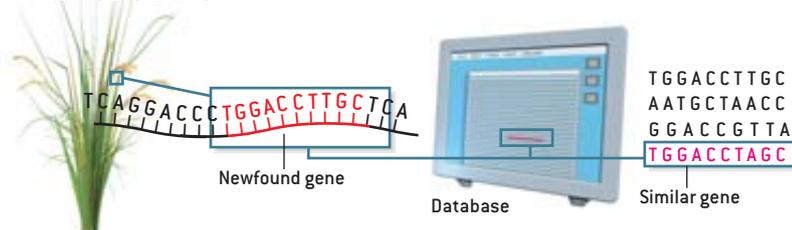
Starch
Proteins
Lipids
Vitamins



MATCHING TRAITS TO GENES

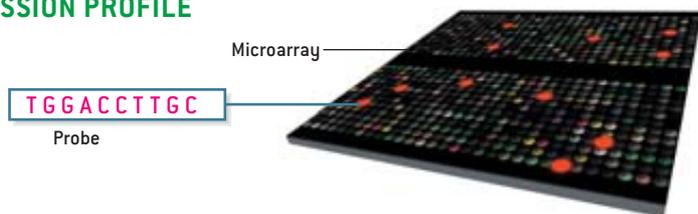
The same tools that allow scientists to trace some human diseases to individual genes make it possible to find the genes responsible for plant attributes. Mapping techniques can narrow the trait-controlling gene's probable location to one region on a chromosome; sequencing of the DNA in that region will then narrow the search to a likely gene. To find out the gene's function, investigators can apply any of the techniques below.

DATABASE SEARCH



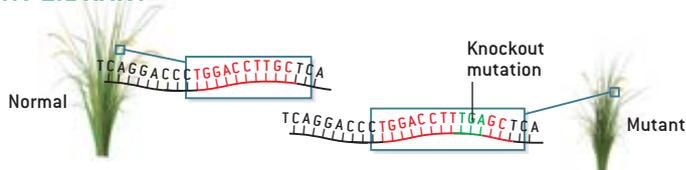
Comparing a newfound gene with known genes in a variety of databases can yield a near match. Of rice's estimated 30,000 to 50,000 genes, 20,000 are similar to genes already studied in other organisms and are assumed to have the same functions.

EXPRESSION PROFILE



Expression profiling gives clues to a gene's function by showing when and where the gene is activated in a plant. A microarray holds thousands of snippets of DNA called probes. Each probe matches a unique signature of gene activity called a messenger RNA (mRNA). When plant cell samples are washed across the microarray, any mRNAs present will stick to their matching probes, causing the probes to emit light. If a gene is expressed, or activated, only during grain development, for example, it is assumed to play a role in that process.

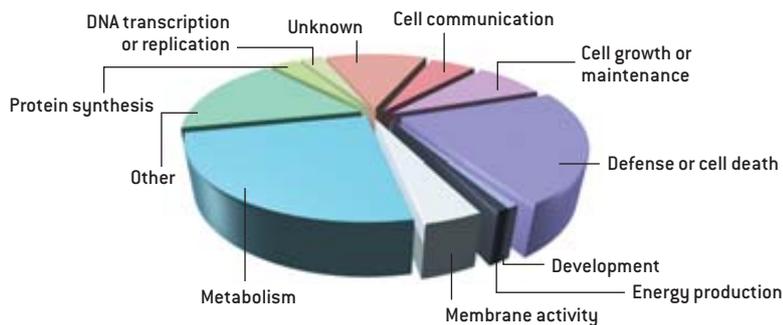
MUTANT LIBRARY



Studying mutants can reveal the function of specific genes by showing what happens when the genes are deactivated. A small piece of DNA inserted into a gene of interest can "knock out," or silence, that gene in the developing plant. Screening the mutant for physical or chemical differences from normal plants can indicate the gene's usual role.

PREDICTED RICE GENE CLASSIFICATIONS

Using the methods described above, investigators have determined or predicted the functions of a large fraction of rice genes.



which allows researchers to predict the role of those genes in rice.

For example, more than 1,000 genes are predicted to be involved in defending rice against pathogens and pests. Likewise, hundreds of genes have been assigned to specific metabolic pathways that lead to the synthesis of vitamins, carbohydrates, lipids, proteins or other nutrients of interest. From experimental data about well-studied plants such as *Arabidopsis* (thale cress), many of the genes that regulate these biosynthetic pathways or affect important stages of crop development, such as flower and seed formation, have also been identified.

A number of research groups have gone further and begun using powerful tools called microarrays to catalogue which genes are expressed, or activated, in a variety of distinct cereal tissues. For example, scientists at our company, Syngenta, examined 21,000 rice genes and identified 269 of them that are preferentially expressed during development of the rice grain, suggesting that these genes play key roles in determining the nutrient composition of the mature grain.

A somewhat different approach to determining a gene's function is to "knock it out" by inserting a mutation into the gene that shuts off its activity and then see what happens to the plant. Sometimes the effect is visible, but the modified plant can also be tested for less obvious changes in any of its normal physiological, developmental, internal regulatory or biochemical functions. Both private and public efforts have completed collections of mutant rice and corn plants in which thousands of specific genes have been knocked out. Such functional genomic studies, combined with sequence comparisons of genes across species, allow scientists to begin developing a basic understanding of how many and which of rice's genes—and by extension those of corn, wheat, sorghum and other cereal crops—contribute to plant development, physiology, metabolism and yield.

Once the function of a specific gene is known, a remaining step in using that knowledge to improve crops is to identify specific alleles of the gene that deliver desirable traits. For example, if a gene is

known to control an aspect of starch accumulation in corn grain, a version of the gene can be sought that functions under severe drought conditions. Such desirable alleles may be found in other modern corn varieties, but even more will probably be discovered in wild relatives of crop plants. Genetic homogeneity among modern crops is an adverse consequence of the way our ancestors initially domesticated them. According to one estimate, modern corn's founding population may have comprised as few as 20 plants. By selecting only a few individual plants with desirable traits to propagate and then inbreeding these for thousands of years, early cultivators severely limited genetic diversity in the domesticated species.

Experimenting with both tomato and rice plants, Steven Tanksley and Susan R. McCouch of Cornell University have pioneered searches for beneficial alleles in wild varieties that might improve modern crops. Their work has demonstrated the genetic diversity available in wild relatives of domesticated plants, at the same time showing that the wild varieties' most valuable resources are not always obvious. In one experiment during the mid-1990s, Tanksley crossed a tiny wild green tomato species from Peru with a somewhat pale red modern processing tomato cultivar. Surprisingly, he found that a gene from the green tomato made the red tomato redder. As it turned out, the green tomato lacked certain genes to complete synthesis of the pigment lycopene, which gives tomatoes their red hue, but it did possess a superior allele for a gene that plays a role earlier in the tomato's lycopene synthesis pathway.

The genetic variety in wild relatives of our modern crops is only beginning to be explored. In rice and tomatoes, an estimated 80 percent of each species' total allelic diversity remains untapped. Remarkable studies by Tanksley, McCouch and others have repeatedly demonstrated the ability of wild alleles to produce dramatic changes in physical aspects of domesticated plants, even though some of the changes seem counter to the wild plants' normal attributes, as in the tomato example. So without the technology to use genes or chromosomal loci as molec-



AUTHORS JOHN SALMERON (left) AND STEPHEN GOFF with experimental corn plants. In the greenhouse, female reproductive parts of the corn plants, called silks, are covered with small white waxed-paper bags to prevent their fertilization by the male reproductive parts, or tassels, at the top of the plant.

ular markers, scientists will find identifying some of these desirable traits or moving them into modern crops nearly impossible.

Marker-Assisted Breeding

ONCE SCIENTISTS HAVE identified specific sets of beneficial alleles from different wild or modern plant varieties, the goal becomes moving just those alleles into a modern crop breeding line, known as an elite cultivar. We could use bacterial DNA or some other delivery vehicle to transfer selected genes, applying the same process (called transformation) used to create so-called genetically modified crops. But scientists are also exploring an approach that avoids the long and expensive regulatory approval process for transgenic plants: breeding guided by genetic markers.

Knowing the exact alleles that confer desirable traits, or even just their chromosomal loci, a breeder could “design” a new plant that combines those traits

with the best qualities of an elite cultivar, then build it through crossbreeding with the help of DNA-fingerprinting technology such as that used to determine paternity or solve forensic questions [see illustration on next page].

All large-scale plant breeding produces tens of thousands of seedlings. But instead of having to plant each of these progeny and wait until they mature to see if a trait has been inherited, a breeder would simply sample a bit of each seedling's DNA and scan its genes for the chosen allele, which serves as a marker for the desirable trait.

Seedlings possessing the desired allele would be grown until they were ready to crossbreed with the elite cultivar. Those progeny would then be tested for the allele, and so on, until the breeder had a population of plants that resembled the original elite cultivar but for the presence in each of a newly acquired allele. The time savings afforded by using genetic fin-

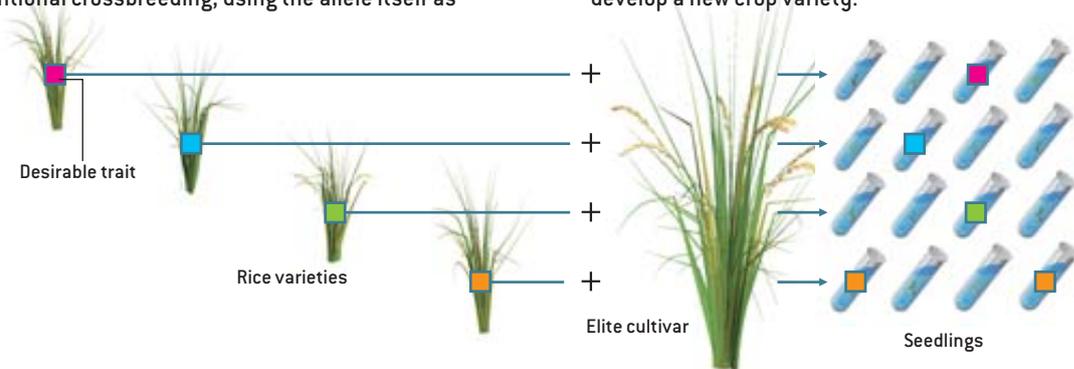
THE AUTHORS

STEPHEN A. GOFF and **JOHN M. SALMERON** are plant geneticists at Syngenta Biotechnology, Inc., in Research Triangle Park, N.C. Goff led a U.S. team in producing a draft sequence of the rice genome published in 2002. He is currently working on a humanitarian initiative to use genetic information about rice to help crop improvement efforts in developing countries. Salmeron, director of applied trait genetics for SBI, has been applying genetics to crop improvement since 1989, when as a postdoctoral researcher at the University of California, Berkeley, he isolated one of the first plant disease-resistance genes in the tomato.

DESIGNING AND BUILDING NEW CROPS

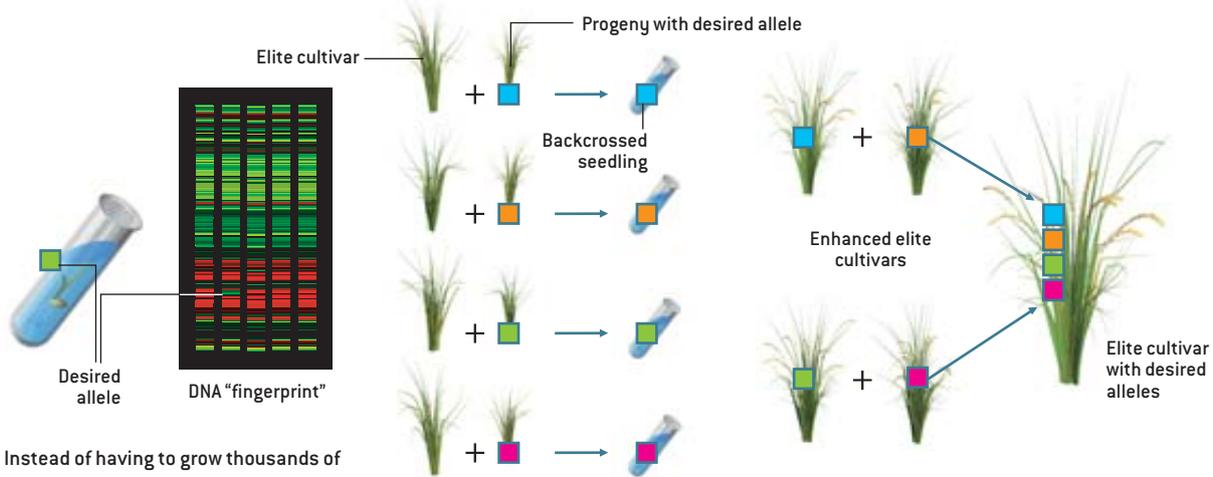
When scientists know which gene controls a specific plant trait, such as seed size, they can search different varieties of the domesticated plant and its wild relatives to find a preferable version, or allele, of the gene. A breeder could then move a desirable allele from one plant into another through conventional crossbreeding, using the allele itself as

a traceable marker for the trait. Instead of waiting a full growing season for plants to mature, the breeder could rapidly find out if seedlings have the desired trait by testing them for the allele in each round of breeding. Such marker-assisted breeding would dramatically shorten the time required to develop a new crop variety.



1 Each of four different rice varieties with a desirable trait can be crossed with an elite breeding line, or cultivar, to produce tens of thousands of seedlings.

2 Some, but not all, of the seedlings will inherit the desirable allele.



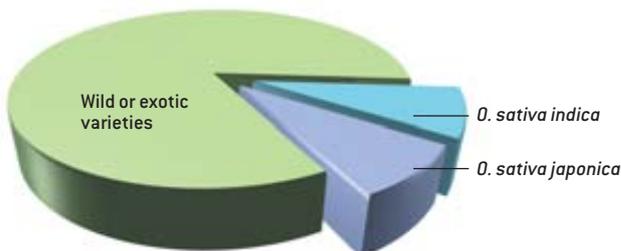
3 Instead of having to grow thousands of plants to maturity to see which ones inherited the trait, breeders can test each seedling's DNA for the desired allele just days after germination with the technology used for so-called DNA fingerprinting.

4 Only progeny with the desired alleles are grown until they are mature enough to breed with the elite cultivar, a step known as backcrossing.

5 Crossing and backcrossing are repeated, with the progeny's genes tested in every round, until all the desired alleles have been moved into the elite crop plant.

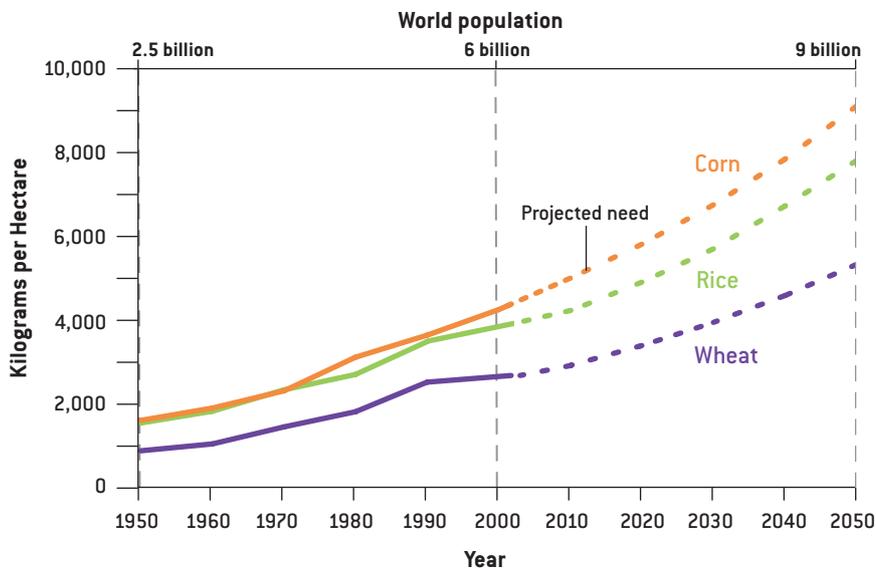
GENETIC DIVERSITY IN RICE

After thousands of years of inbreeding, modern crop varieties are far less genetically diverse than their wild relatives (*pie chart*), making the wild plants a rich reservoir for novel alleles. The untapped wealth in wild plants is not always obvious: in experiments with rice ancestor *Oryza rufipogon* (left), alleles from the wild plant were moved into a modern high-yield Chinese rice variety (right) using marker-assisted breeding. The low-yield wild plant's genes raised the modern variety's yield by 17 to 18 percent.



SLIM FILMS (illustrations): RICE PHOTOGRAPHS REPRINTED WITH PERMISSION FROM TANKSLEY AND MCCOUCH IN SCIENCE, VOL. 277, 1997. © 1997 AAAS; C. MARTINEZ/International Center for Tropical Agriculture (left); CORNELL UNIVERSITY (right)

CEREAL CROP YIELDS



WORLDWIDE AVERAGE YIELDS for corn, rice and wheat nearly tripled between 1950 and 2000, a period that saw global population do the same. To feed a projected world population of nine billion in the year 2050, while maintaining present average consumption of 0.9 to 3.3 pounds per person a day of these cereal crops, yields must keep rising by 1.5 percent a year.

gerprinting to test for trait markers at each stage of this process would literally shave years off the time typically required to develop new crop varieties. This acceleration would allow breeders to become more responsive to changing circumstances, such as the emergence of new pests or of resistance among old pests to current countermeasures. Tailoring new crop varieties with combinations of characteristics that are optimized for different environments, farmers' needs or consumer preferences would also become easier.

But the real revolutionary potential in this method lies in its power to open up the genetic bottleneck created thousands of years ago when our major crops were first domesticated. Once scientists accumulate more information about the functions of genes in the grasses, we can more effectively search the huge reservoir of genetic diversity waiting in the wild relatives of modern crop plants. One of McCouch's experiments provides an example of the possibilities: she used molecular markers to identify gene loci in a Malaysian wild ancestor of rice known as *Oryza rufipogon*. McCouch and her colleagues then employed marker-assisted breeding to move a total of 2,000 genes—approximately 5 percent of the rice genome—into

plants of a modern Chinese hybrid rice variety.

This experiment was focused on finding alleles to increase the already high-yield hybrid's output still further, and the resulting test plants were examined for several yield-improving traits, such as plant height, length of its flowering head (panicle), and grain weight. Approximately half the wild relative's loci turned out to have yield-improving alleles, although some of these also had negative effects on other aspects of the plants' growth, such as slowing maturation time. But two of the alleles from *O. rufipogon* seemed to have no negative effects and produced yield increases of 17 and 18 percent, respectively, in the modern cultivar. As in Tanksley's tomato experiment, nothing about the wild plant's appearance [see box on opposite page] suggested that it could teach modern rice some-

thing about yield, yet the results were impressive and encouraging.

Of course, certain beneficial genes cannot be moved into modern crop varieties by means of traditional breeding. For example, genes conferring some types of herbicide tolerance or insect resistance do not exist in plants that will crossbreed with corn.

A gene can be transferred into a recipient plant using current transformation techniques, but they do not allow scientists to specify where in the recipient organism's genome a new gene is inserted. Thus, one could add a new allele but not necessarily succeed in replacing the old, less desirable allele. Yet in the cells of mice and in some microbes, a phenomenon called homologous recombination directs an introduced gene to a chromosomal location whose DNA sequence is most similar to it—permitting a desirable allele of a gene to directly replace the original.

In the future, we may be able to achieve the same one-step substitution of gene alleles in crop plants. Homologous recombination was recently demonstrated in rice, and a related process has been used to replace alleles in corn. Once it becomes routine, the ability to swap pieces of chromosomes this way in the laboratory might allow scientists to exchange alleles between some plants that cannot be crossbred naturally.

Today marker-assisted breeding is already speeding the process in crops of the same species or close relatives. No new cereals have been domesticated in more than 3,000 years, suggesting that we will need to rely on improving our current major crops to meet ever increasing food demands. By providing the ability to peer into plants' genomes and the tools to harvest their hidden treasures, genetic science is opening the way to a new green revolution. SA

MORE TO EXPLORE

Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. Steven D. Tanksley and Susan R. McCouch in *Science*, Vol. 277, pages 1063–1066; August 22, 1997.

A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *japonica*). Stephen A. Goff et al. in *Science*, Vol. 296, pages 92–100; April 5, 2002.

Gramene, a Tool for Grass Genomics. D. H. Ware et al. in *Plant Physiology*, Vol. 130, No. 4, pages 1606–1613; December 2002.

Early Allelic Selection in Maize as Revealed by Ancient DNA. Viviane Jaenicke-Després et al. in *Science*, Vol. 302, pages 1206–1208; November 14, 2003.