

October 15, 2004 – 10:15 p.m.

STRATEGIES FOR ACHIEVING A DOUBLY GREEN REVOLUTION IN AGRICULTURE

Dr. Susan McCouch

Professor of Plant Breeding and Plant Biology
Cornell University

First, I'd like to thank Dr. Brady for that nice introduction and thank Judith Pim and Ambassador Quinn and all the organizers of the World Food Prize for hosting such a wonderful event for bringing a lot of people together at a time when sometimes I think we should be focusing on what we can do rather than on the dismays of the current situation.

Today I'll talk to you about the discovery and transfer of trait-enhancing QTLs, and I don't want to bash you with that acronym. That stands for "quantitative trait loci," and instead of calling them genes, we call them QTLs because in many cases there are multiple genes that confer a trait or a phenotype, and we are identifying them in terms of a region in the genome.

I can skip quickly over these first slides, but the reason I left them in is because it's an argument that we have to use when we talk about why we have a rice breeder at Cornell University (Please refer to power point). It's an important point. Cornell obviously doesn't jump to the front of your mind as a point of major rice production either in the United States or except under our greenhouse. And the second thing that I want to mention is that Cornell had never had a rice breeder before. So the fact is that the importance of rice and the opportunities that the rice genome has opened up to us has given us an opportunity for people who are in non-rice growing parts of the world to join the great team of international scientists working on this crop.

The world population growth curve is simply a startling one and one to keep in mind as we go forward. I think it was Dr. Borlaug who said, "If you don't keep the pipeline full technically or technologically, then all of the policy decisions won't help." So what I've focused a lot of my work on is trying to keep the technology pipeline full, so that when we have the appropriate opportunities for distribution and the discussions about how to foment opportunities among small farmers, there will be in fact ideas about what they might be growing.

I'll talk about rice as a member of the cereal family. The cereals together represent 60% of world food production, and rice shares a common set of genes with its sister grains, as Dr. Swaminathan has just mentioned. Rice is the first crop species to be fully sequenced. That's an accomplishment, and one that has given us an opportunity to look again at what the impact of genome science will be on crop improvement.

What will the knowledge about genes and DNA sequence actually contribute to a new “Doubly Green Revolution?” I think we all know that genomics, or the genetic potential of a plant, is a critical feature of what we’ve come here to celebrate with the World Food Prize this year. But what is not obvious is how knowing more about the genes or the DNA sequence is actually going to change the way we think. I hope I’ll be able to give you some insight into some of the kinds of changes that I think are coming.

First of all, just keep in mind, genotype, phenotype is a relationship. The genotype confers the genetic potential, but the potential of a genotype is only realized in the phenotype under the appropriate environmental conditions. On the genome side, we’re interested in looking at the DNA, the RNA and the protein because all of these things work together in very complex interactions to give rise to the phenotype. So far today we’ve been talking mostly about genetic gains based on the identification of a superior phenotype in an environment which is managed in terms of inputs and plant spacing and other things.

What I’ll be talking about is – What is going on under the hood that helps us understand when we have a superior genotype? I think when Ron Cantrell comes to speak, he’ll be talking more about the entire package that’s going to be needed to deliver these new advances. But what is intriguing, I think, to scientists today, and hopefully brings in a large number of new scientists in the future, is to use an example of a plant that has the potential to feed so many people in the world and yet look at it in the scientific arena simply for curiosity’s sake and then try to link the two. In other words, the genetic understanding of what’s going on in a plant, link it to the yield potential and deliver that as a knowledge-based set of tools that the breeders can use to produce new varieties.

Conceptually, this series of funnels represents the phases of the breeding process (Please refer to power point). All modern varieties are in fact derived from the wild ancestors, and the wild ancestors had a large amount of genetic diversity, much of which has been eliminated or left behind over the process of continued plant improvement. When we talk about early domesticates, we’re talking about land rice varieties, varieties that were selected by our earliest agricultural ancestors. They were simply the plants that performed best in the areas in which the humans were selecting them, but also they had the features that human beings prefer – larger grains, nonshattering types; sometimes they were simply sweeter or more tasty or more aromatic.

The little colored lines here that outside of these funnels represent the alleles that were left behind as we continued in our selective process up until the modern times when we have a much narrower genetic base. But we have selected for things, for genes and alleles, that allow our varieties to do very well under the high-input conditions of modern agriculture. These alleles that were left behind are not necessarily bad. They were at the time the things that were discarded because of the superior phenotypes of the plants that were carried forward.

The hypothesis on the table is to look at how genetic diversity, the utilization of genetic diversity and the combinations of diverse types, such as you see in the hybrid rice production that Dr. Yuan Longping spoke about, how combining novel or diverse genetic or alleles of genes can actually augment performance even though the phenotype, the original phenotype was not superior. I’m going to give you an example.

So the focus of this research is – Can we identify genes and quantitative trait loci (QTL) from wild and exotic germ plasm that will improve the performance of our best cultivars. The best cultivars are high-yielding, well-adapted and contain many of the superior features that we want as we move ahead. The wild and the exotic ancestors have very few of those features when taken at the pheno-typic level.

They often look like something you would not even want to plant in your field, and here in the United States we're forbidden from planting the wild relatives of rice anywhere in the United States. These are considered to be noxious weeds. So the idea that you could harvest valuable genes from these wild species requires some level of genetic insight, since it is not a common practice in a plant breeding program oriented towards rapid genetic gain.

But the fact is, when you cross the wild relative, which is kind of spindly looking, with one of the best cultivars still widely planted today, the F-1 shows immediate heterosis, even in the young vegetative stage. Every one of the crosses we make to our modern cultivars with our wild ancestors shows heterotic vigor, or hybrid vigor, in the F-1. The question is – Can you take that vigor that you see in the seedling stage and translate that into yield or seed production in a cereal?

We undertook a global experiment. It's a largely collaborative effort and involves many people in national programs and international programs throughout the world. We asked them to select the best of the cultivars that they were growing at the time and cross it back to one of the wild strains, either from Asia or from Africa, to see whether introducing genes that have not been inherently part of the cultivated gene pool might actually augment performance.

I won't go through this in great detail, but there was crossing that was done by the national program scientists, multi-location field testing, massive pheno-typic evaluation; and then our lab inherited a student from each of the programs that had undertaken the crossing. We did the marker analysis on the remnant seed from the same populations that were being grown throughout the world, these interspecific progenies that were a result from two generations of backcrossing. We did QTL analysis where we pulled together genotype and phenotype information, purified some of those lines, and those lines were then sent back for multi-location testing in the locations where the recurrent parent had originated. Some of them have gone on to produce improved varieties.

At the same time in our lab, we used those targeted lines called, "neoisogenic lines" containing individual QTLs for the gene cloning or gene discovery process in order to provide new selection tools that could be used to help people in this selection process in the future. You can see that this requires a large collaborative effort, and it would not have been possible without all of the contributions from the national program scientists who've been such strong collaborators throughout the last eight years of this program.

We undertook a standard evaluation on these 12 characteristics that are the standard agronomic traits associated with rice performance, and many of them undertook additional ones as well. The fact that we had a standard set of final types allowed us to do some comparative work as the data came in.

This gives you an idea of the range of environments we were testing in. We had trials in Korea, in Chung Sha with Dr. Longping Yuan's group, in Indonesia with Monty Jones at WARDA, in the southern part of the United States, in Colombia and in Brazil. Each one of these was a different recurrent parent selected for adaptation to the environment. Some of these were upland environments, some of them were lowland, temperate, tropical, etc. They all used the same wild donor in the crossing program.

The genotyping was done using electrical markers that had been developed in my lab that are called "simple sequence repeats." Just to give you an idea, we now have over 2,500 of these that we have been using globally, and they align us to both the physical and the sequence-based maps. When you put together molecular marker information and phenotype information, you gain incredible insight as to which genes on which parts of chromosomes are contributing to your trait.

This is just to show you a histogram of a yield profile, grains yield per plant. The wild parent, of course, is very low-yielding. The recurrent parent in this case, 23 in Korea, was quite high-yielding, but what you see here is a portion. In every one of the histograms done by every one of our collaborators, there was always a proportion of plants coming out of these BC2/F2 families that performed better than the better parent. And the question was – Could we understand the genetic components that contributed to that better performance? And if we could understand actually which genes were involved, could we use that information to make predictions about how to do this again in the future?

When you combine genotype and phenotype information, you're able to identify a region along a chromosome, in this case on chromosome 1, associated with the yield potential of this proportion of the offspring.

This shows you the region. It's marked with one of our molecular markers, and you can use that molecular marker to ask a question about all of the other populations that have been evaluated in different environments and in different genetic backgrounds and ask whether any of the same introgressions contribute to increased yield in those different backgrounds.

So the first step is to ask where all of these yield-enhancing genes, earliness genes, disease resistance genes or grain quality genes are located. If you look to see where the red markers are, in other words, the grains per plant, these phenotypes, and you ask, what is the component of the yield that is giving rise that is explaining that yield enhancement, you can get yield enhancement from either increasing the number of grains per plant or increasing the thousand-grain weight per seed (Please refer to power point). Obviously, in rice improvement we're not interested in increasing the seed size tremendously – we'd like to simply produce more grains per plant.

In the chromosome 1 introgression, we see a yield advantage associated with an increase in grains per plant. We see the same thing in this region on chromosome 2, and interestingly there is a region on chromosome 3, just to show you the alternative, where the increase in yield is associated with an increase in seed size. There's another over here where it's increasing corresponding to grains per plant. That tells you that there are a nutrient of regions of the genome where we would like to identify offspring that have the introgression from the parent but not any

of the other introgressions throughout the rest of the genome. We do that through backcrossing. It does not involve genetic engineering because we still don't know at this point what the identity of the genes are. We just know the region, and we have a molecular marker.

As the results came in from our different collaborators – we have students from each of these countries doing the work between the field and the lab. In an upland environment in Brazil we see it even in one of the high-yielding inbred lines, tropical japonica lines from the southern part of the United States. This was evaluated in Colombia, because we weren't allowed to evaluate this in the United States because of the fear of the wild parent. These lines are now being evaluated in the U.S., and they're showing yield advantages due to this introgression.

So this is the kind of data that came in from this collaboration, and it gives us now a very focused arena, a smaller area of the rice genome, in which to look for the gene that is conferring this yield advantage in these different genetic backgrounds and these different environments.

I just want to show you then Dr. Yuan's recent super hybrid rice release. He in fact took advantage of one of these yield introgressions. The proof is down here at the molecular level, and it's due to a number of things. But he has always claimed that this introgression, which he has held onto, using molecular markers, is part of his yield advance in this new hybrid. So he's comparing the panicle. You see it carries many more grains per panicle, and it's going to be very similar to the story that Monty Jones told. And I'll come to that in a moment.

So the question for molecular biologists and in fact I think for many intellectually alive people is – well, that's pretty interesting, so now we can do something with marker-assisted breeding; we can move this region of the genome through backcrossing into any one of a number of our high-yielding varieties. But what are the genes that underlie this? There is a region of chromosome, but we'd love to know what these wild QTLs are actually doing and how they work.

So the problem is, you have to clone those genes. And when people think about biotechnology, often they think about transgenics. I don't know how many of you have stopped to think about the fact that there really are not a large number of all possible genes that are available for transgenesis. In other words, we really don't know much about what genes do. We have their sequence, but it's an ongoing quest to try to understand how they all work together to confer the kinds of phenotypes that are important to produce the food that we all depend on.

So we have to go about this. We call this the “needle in a haystack problem.” The rice genome has 50,000 genes. We'd like to know – we've narrowed down the search space; we know it's a smaller region – but we'd like to know which of those genes is responsible, for instance, for that yield advantage I just spoke about.

So you can use molecular maps and markers to do that, and I think Dr. Tanksley will be giving you some examples of how that's done. You can also use sequence information if you have an idea of what gene it might be, from known pathways, research in other organisms – that can give you a pretty good idea, and you can test those hypotheses. We can use the kinds of comparative maps that Dr. Swaminathan just spoke about in cross-species analysis to in fact,

once we do know what genes confer the advantage in one species, to predict what advantage they might have in another species.

So the power of this genome research is that it can take you beyond the specific example at hand and produce information that can be used for predictive purposes.

This is just a very brief discussion about map-based cloning. This was an example of work done in my lab. If you use positional cloning, you can identify almost any gene responsible for a trait if you can measure the trait reliably. In our case, we went after the semi-dwarf gene, the Green Revolution gene as its known, and we know that it resulted from a knockout mutation which deleted 132 base pairs in a specific region of the GA20 oxidase gene. That was published simultaneously by about three or four laboratories in 2003.

What's interesting is that mutations in the same GA20 oxidase gene were also known to confer semi-dwarfism. So here's a tall plant, here's a semi-dwarf, and here's a dwarf. So it's been well known for many years that this particular set of genes had an association with dwarfism, but nobody knew which of the many genes associated with dwarfism might have given rise to the Green Revolution fina type that was so valuable. In fact, as Dr. Khush well knows, and I learned from him, there are about 65 known dwarfing genes in rice whose chromosomal positions are well known, some of which have been cloned.

So the problem is a quantitative one. There are many genes involved with almost every trait that's important in agriculture. In this case we had a semi-dwarf habit. It was also associated epistatically with a difference in the panicle size. The question was in this developmental pathway of genes that had been identified in Arabidopsis, which one conferred the semi-dwarf trait in rice? And perhaps of interest to Dr. Borlaug, the question might have been – Was the semi-dwarf gene that was responsible for the Green Revolution in rice, the same gene that was responsible for it in wheat? After all, the phenotype is remarkably similar.

Well, here's the pathway involved in what's known from Arabidopsis, and here we show you that the semi-dwarf gene in rice is called GA5. That's Arabidopsis terminology. But it turns out that when this experiment was done, the gene that confers the Green Revolution semi-dwarf fina type in wheat was simply a different gene in the same pathway. It's called the GAI gene. The difference is that GAI is in fact gibberelic acid responsive, while the GA5 gene was not. It's also quite interesting, if you look carefully, that there are two of these GA5 genes in the rice genome, and they are differentially expressed, and had you knocked out the other one, you would not have gotten the favorable phenotype.

We can then use our comparative maps to make predictions about where each of these genes lies. Now that we have rice genome sequence, we know where they lie in the rice genome, which is this inner circle, and we could tell you with a high level of predictive power, where the other genes and the other grass species are likely to lie.

So this is a very powerful kind of information. It takes you from the specific example, and it allows you to make inferences about other species and other genes and other pathways that you might not have yet discovered in a plant breeding context.

These then can be used for marker-assisted selection. This is just a very simple profile showing you some things that might in fact surprise you. If you wanted to use marker-assisted selection to find out whether the specific allele had been inherited, you might be surprised to find that sometimes a phenotype is not predictive of the genotype. It's not surprising to a plant breeder, but it's always one of the things that I think is helpful to think about.

This has the fully competent gene at this GA20 oxidase locus. The IR64 IRA and Lamont phenotypes are all semi-dwarf; they all inherited the knockout gene (Please refer to power point). It's 132 base pairs smaller. However, interestingly, neponbere, which is a semi-dwarf, high-yielding phenotype, is semi-dwarfed for a different reason. It's not because of this particular gene in this pathway but perhaps one of the many others that we know affect stature.

Interestingly, again, we see that many of the previously characterized mutants, many of which were not favorable in agricultural terms, are turning out to be very, very useful tools as we try to clone genes underlying quantitative trait loci. Because they in fact tell us about which gene is responsible for a QTL, although the increased information comes when you clone it using naturally useful sources of genetic diversity, is that the specific allele that is valuable to a plant breeder may not be represented in your collection of mutants. So the collections of mutants are a valuable starting point for identifying essential genetic loci that might be contributing either to the phenotypes specifically, or at least genes involved in a pathway that open up the door for future inferences about which genes are involved.

Something happened. This is my picture of Monty in the field. Monty, I'm sorry. I think it's because you've already given us so much good information. This was a picture of Monty in the field, and we can't recognize him here, but fortunately we have him in person. And this is, of course, Monty's work. I just want to say that the new rices for Africa are, in fact, the transgressive segregants, they're inbred lines. They're transgressive as he has explained to us. They're derived from the same kinds of intraspecific crosses that I've been working on, and he and I are collaborating. We're very interested to know whether we can now get down to not only the genes that confer branching – and many of these genes have already been cloned in maize and we'll be very interested to see whether we can understand which genes have differentiated in the pathways and then further go on to see whether any of our candidates for these yield genes that enhance seed number per panicle are in fact similar in his sativa crosses.

I will finish up by saying that I think that new gene-based strategies for crop improvement are in fact essential to meeting the nutrition and productivity requirements of the future. They will contribute not only to productivity gains but to some of the other shifts that we're hoping to make in terms of nutritional aspects, specific biotic and abiotic stresses. The power of these technologies is that they don't have to go after only simply inherited traits, but the quantitative traits are now opening up. They are a tremendous playground for those of us who enjoy understanding how genes work and making sure that the genes that we are targeting for our work ultimately are important to the production of the future.

I want to acknowledge not only the students from Ivory Coast, Indonesia, from Colombia, from China and from the U.S. who have been involved in this wild QTL project – Lee is from Dr. Longping Yuan's program, and Monty Jones' program. For the high school students who may still be in the audience, that we've had high school interns every summer in my lab,

and in 2002 this pair one the first place in the Westinghouse Science competition in the United States – and we're very proud of them, partly because it's the first time that an agricultural project has ever one first place in the Westinghouse Science Fair.

I also want to thank my national program and international collaborators, without whom none of this would have been possible. Thank you.