

October 15, 2004 – 10:15 p.m.

STRATEGIES FOR ACHIEVING A DOUBLY GREEN REVOLUTION IN AGRICULTURE

Dr. Steven Tanksley

Professor of Plant Breeding and Plant Biology
Cornell University

First, I want to thank the organizers for inviting me to participate in this important conference, such a prestigious group of people. I have to confess I was a bit surprised when I was invited to speak, because it's been some years since I did any work on rice. So I figured it was either a mistake on the organizer's part or, given that biodiversity was part of the topic, they brought me in to increase the diversity. But for whatever reason, thank you, and now you're stuck with me but for a very short period of time.

First, I want to make a comment, and that is that, whether we think about it deliberately or not, we have entered into an amazing scientific revolution, and it's a revolution not just of knowledge but of understanding. And all of us, I think, have a sense of this, whether we can articulate it or not, and this revolution is manifest in many ways, one of which we all think about, such as secrets in the human genome or the rice genome or the some other 150 genomes that have now been sequenced.

But the result of this revolution will be beyond any scientific revolution that's occurred before. In the last century, we had the great period of discoveries in physics, which revolutionized part of our thinking about the universe. We've had other revolutions, for example, in flight and the Apollo program that took place in the mid-part of the last century.

But this revolution is about the most fundamental thing that there is, and that is life. And this is what gives us interest in it, because it's about us and we are alive. It is the underpinning of all of our economy, or will be, the biological basis for everything from plants and food to nutraceuticals, pharmaceuticals, and in the future the way things are manufactured and what they're manufactured from.

So when great scientific revolutions occur, and this one will occur within our lifetime, it raises great possibilities. But in order to take advantage of these possibilities, we have to imagine what it is that we would like to discover. And one of the most restricting things that can happen in science is, you begin to believe that what you can discover is only what is technically feasible. And in the past, biology has been very limited by what was technically feasible, so we often craft the questions on the basis of technical feasibility. But now with the advent of genomics and all of

the other technologies that will allow us to understand the blueprints of life, we have the possibility to understand virtually anything we wish to go after.

And it turns out that one of the most, perhaps the most, fundamental question about life and biology, which has gone unanswered since the beginnings of the Darwinian Revolution, is also the most fundamental question about how we improve and feed ourselves with regard to plants. And that is natural variation. Now, this is something we all have a sense for, too. We can look around the world at the diversity among us, and we know there are differences and that there are genetic bases for these. We also know, as we go out and walk in nature, a huge diversity within a species and between species.

But the reality is – we don't know what is causing that diversity. We've created the Green Revolution. We continue to improve crops and other organisms to our benefit, largely in absence of knowing what the basis of that variation is. Likewise, all of medicine, as advanced as it is, has taken place with little or no knowledge about what the basis of differences are between humans that makes them susceptible to different diseases.

So at this moment in time, the most fundamental question about agricultural science is now the most fundamental question of plant improvement. And that is – what is the basis of natural diversity, and how does that knowledge relate to how we continue to feed and clothe ourselves. So I want to continue on in the vein that Dr. McCouch presented in her talk, which is utilizing natural diversity. And I think she pointed out very nicely that, within the natural diversity and the wild relatives of our crop plants, exist varieties of genes which can tremendously improve the productivity, resistance and tolerance of crop plants, but this variation has largely been hidden by looking strictly for fine types. And she showed very nicely, bringing out a gene from the arufa pogon could increase the yields of rice, not in just one location but multiple locations.

So if we think about what she presented, which is a strategy to go in and harvest genetic variation from our wild ancestors of crop plants and make great improvements for which there is great potential, we should also ask – what is unique about that variation versus any other way we can improve plants? In other words, what is the pool of genetic variation that we're going to try to harvest?

If you look at the genetic maps created after the Mendelian Revolution up until about 15 years ago. If you look what's on those maps, you'll be surprised what the genes are. So, for example, if you look at the plant maps, you'll see – and this is true of maize or tomato or rice or any of the human maps developed prior to this recent era of molecular biology – there were genes such as genes for dwarfism, genes for sterility, genes for lack of anthecyanin, genes for lethality even.

If you look at the human maps, it's even more startling. Those maps are comprised of genes for mental retardation, for dwarfism, for lack of limbs, for being albino. It seems very unlikely in the divine wisdom of the universe that we would have a genome full of genes like that. The reality is, when we've gone back now and begin to study the genes that gave rise to genetics in plants and animals, we've discovered that most of those are mutations for loss of function.

So the left side there, which is, if you look at the yellow is a chromosome and the red is the genes along it, if we mutagenize a gene and knock the gene out or cause it to be dysfunctional (it has a white X through it), we get a dramatic final type, such as a plant that grows very slowly or very weakly. And so we get a nice 3 to 1 segregation ratio. If we mutagenize another gene next to it and knock it out, we may get an even weaker plant. But at some point in time, if you carry this on to its inevitable end, as we thought, lose genes, we eventually become lethal because genes are not easily gotten rid of – they're there for a reason. So if we think about – is this the way that evolution proceeded or plant breeding is going to proceed, the answer is probably no, because it's an evolutionary dead end.

Now, if we look at natural variation, it has a very different set of properties. One is that the differences between plant varieties for yield or virtually anything you want to look at, or the differences among us in this room, are largely not single-gene differences – they're polygenic. And being polygenic, they've been difficult to study with the tools of Mendelian genetics and molecular biology, because they don't give discrete classes of segregation.

The third thing – and this is extremely important – the variation that differentiates plant varieties, that differentiates species, that differentiates each of us from each other, are not differences that occurred yesterday. They didn't occur in the lab because someone mutagenized us. These are varieties that have been passed down through generations, for thousands, tens of thousands and sometimes millions of years. They've been passed through the process of selection. And it's very unlikely that those changes are just knocking out genes, because again this would be a lethal approach to evolution in breeding.

So adaptive evolution – and this includes plant breeding – gives rise to something very different, not a weak or a mutant or a dwarf of something that's lethal, but to changes in appearance or function.

So on the bottom here I show three nightshade plants, which are in the family Solanaceae, the tomato, eggplant and pepper. They have the same basic genome, 12 chromosomes, similar to the grasses, Poaceae, and basically a gene-for-gene relationship. But anyone can tell the difference between those three species. The question is – how did it get to be different? What actually happened?

So we started thinking about this question. And you have to focus somewhere – you can't do everything all at once. So we began to ask this question in a very confined situation, and that is one species tomato, and about a very specific trait, and that is fruit. And the reason we chose this is that, like most plants are domesticated, domestication often did something quite extraordinary to the plant. In the case of maize, it gave rise to the large ear, which is not found in the wild ancestor, teosinte. In the case of tomato, the large Beefsteak tomato we see in our supermarkets, is really not what's seen in nature. A true wild tomato, which still grows in South America, is a very small berry, about the size of your little finger. And these are completely cross-fertile. You can make F1s, F2s, any generation you want, but you don't get a 3 to 1 segregation ratio; you get a continuous segregation.

So we began to ask the question – what actually happened to allow plants to evolve fruit? And this is also a very classic quantitative trait that had been published from the 1930s on – the fact that fruit size and shape was highly quantitative character.

So the reason that no one had been able to get at the basis of what caused these changes from the small wild to the cultivated fruit was that, again, it was monogenic, polygenic; so you couldn't fit it into the Mendelian model, nor could you fit it in molecular biology – you've got these continuous distributions.

So the quantitative trait mapping that's come about in the last 15 or so years and that Dr. McCouch talked about, we now have a tool to go in to look at any two individuals, any two species that differ for any characteristic we can measure. If we can get sexual reproduction from those in any form, we can deduce the positions of the chromosomal locations which are causal. It doesn't show us what the genes is, but it tells us where in the genome they are.

So we've done this for a large number of crosses between wild and cultivated tomatoes and between various cultivated tomatoes. And the surprising thing is, if you look at all the variation for size and shape found in all the different types of tomatoes, it really distills down to about 30 loci that were affected, and that's out of about 35,000 genes that make up a tomato genome. About five of those have larger impacts, and the other ones are more minor impacts.

So first of all, it said that there is a manageable number of genes involved in this process to understand. So we set about cloning these, and we've cloned a number of them now. And I'm not going to go into these in any detail because there is no need of it, but I want to focus on what types of genes have been affected and what have been the effects on those genes that gave rise to this useful variation – and also some of the surprising lessons we've learned as we've gone along the way.

So the first one we cloned was a Quantitative Trait Loci (QTL) for fruit size that watershed one of the more significant ones and also one that occurred very early in the domestication of tomatoes. It was one of the early steps from which the small berries began to get larger through human selection. And this is a QTL on chromosome 2. We fine-mapped it down to a piece of DNA small enough to sequence. And we knew genetically this piece of DNA had to have the cause of the QTL on it, but also we had no idea what a QTL was.

So we sequenced it. We found a single gene on it, called CDA44. And this was a gene that was perfect. It was one that was expressed throughout fruit development and was located in the exact right place, and we had lots of cDNAs for it, so we assumed it was the gene. We transformed it in all manner you can imagine, from expression, knockout, etc. – and we had no effect on the fruit at all, which is very disconcerting for me partially but especially for the students that are trying to make careers out of this. So I had to keep them encouraged while also having great sense of strain myself. But we finally had to realize that either this was not the QTL or it was something so extraordinary that we couldn't get at it with the typical tools of molecular biology.

But one thing that we did do – and this has become one of the hallmarks of genomics – is to use informatic technology to search for things. We used a lot of other multiple searching

programs to see if there was any other possible genes in that piece of DNA. The first programs we used just found the one gene, and also the one gene was the only one that had any CDNAs for it, any evidence of expression.

But one of the gene-finding programs found a very small open reading frame which we referred to as the 4th X or unknown. The problem was it had no CDNA. We had screened a million primary plaques during fruit development and never gotten one. And to this day, even though there are 200,000 ESTs for tomato in the genomes mean sequence, there's never been a CDNA isolated for this particular gene. So we didn't know if it really was a gene. We could do RTPCR and show the introns are spliced out, so it had to be a gene, although expressed at very low levels since there's no CDNA. When we transformed it, it was the QTL. So this was the first surprise, that a QTL could be a gene expressed at such low levels that it's virtually invisible, yet it had such a profound effect on the plant.

The next thing that was a big surprise to us – and this shows how little you may extrapolate into the unknown and cause great trouble – is we had imagined getting larger fruit would be an active process. You would be somehow overregulating something. Turns out it's just the opposite. This QTL, to make large fruit, is actually a QTL to stop cell division. So these small berries, which have to be eaten by birds, have to stop growing at the right size to be small enough for the birds to carry off. Well, what had happened in this case, a mutation had somehow changed that, so this gene was not controlling cell division as it should. Therefore, too many cells divided and the fruit got “too big” from the plant's viewpoint but good from our viewpoint. So it turned out to be a negative regulator of cell division, which was another big surprise.

We then began looking at what I think is the most important issue, and that is – what had actually changed genetically? We had the gene in hand, and from the gene you can deduce the protein. And all of our thinking about genetics has been based on genes' coding for proteins, proteins causing evolution. So we began looking at this protein to ask what could you do to the protein to make it change its function. And the two alleles we had cloned, the large and small fruit, had had three amino acid differences in the interminants. So all of those became candidates.

We used a form of association genetics, which is a simple way of saying we looked at a lot of different tomatoes with large and small fruit alleles, sequenced them all, and asked – did any amino acid have coincidence with large and small fruit? And the answer was no. In fact, we found some small fruited types that had exactly the same amino acid sequence as the large fruited type.

So this said that this protein essentially didn't evolve at all. But yet the protein, the gene coding for it, was the cause of how they got domesticated tomatoes, which is a bit of a puzzle. So I began thinking back. I was a graduate student at UC Davis during sort of a peak period I'd say of evolutionary biology and the discovery of molecular biology, which was enzymes back then. And when people found enzymes had variation in natural populations, they assumed that was the basis of evolution and devised all types of theories.

But an evolutionary biologist from Japan had a different idea, and he has put it in a very nice way. His idea was that you don't evolve proteins very rapidly, but you evolve how they're regulated – and that's what makes all the difference in evolution. But his parable was simpler. He

said, “If you give me the coding protein regions of a chimpanzee and the regulatory regions of an elephant, I’ll give you back an elephant.”

So we began looking for the elephant. We sequenced upstream in the promoter to see if there were any differences. And we found that there were eight mutations in the promoter that were always present or absent if you got large or small fruited tomatoes. So this said there was something in the regulatory region that was consistent with regulation, but the surprise was there were eight things that had changed, not just one; and we didn’t know which, if any of those, were causal.

We also didn’t know at this time if it really was regulation, because this could be coincidence, so we had to do very controlled developmental studies on a gene that’s hardly expressed at all, in a very short period of time. So it was a very demanding study, but as replicated in the field in greenhouse and using RTPCR to detect this very rare transcript, what we found was the most significant difference that caused the change in the fruit size was that the protein, the gene had changed expression in its timing by about seven days. In other words, the fruit that stays small, the gene comes on gradually just after anthesis and stays on for a long time during the cell division period and shuts off, after which the cell elongation occurs. And the one that gave rise to the large fruit, the gene came on quickly and disappeared. In other words, it wasn’t doing its job.

So this is somewhat like life where they say in life that in music, sports and in falling in love, timing is everything. In this case, timing was everything. The gene has to be present at the right time, or that protein does, to do the process. And that very minor shift in the very minor gene made all the difference.

So as I said, we’ve now cloned several of these. And I just want to summarize the story for the ones we’ve looked at. And even though this is tomato and may be sacrilege to talk about in a rice conference, there have been a number of other QTLs cloned now in other plants and in animals. And there’s some sort of a picture emerging. I think this represents a bit of the diversity of that picture.

Again, the first one we cloned was a regulatory mutation that changes timing, which is called heterochronic mutation. The second one, which is a QTL-controlling shape, was not a change in the promoter of the gene but a change in the protein itself, which is a regulatory protein – so in that sense, it was a change in regulation. The third one on the right is one that’s, if you can tell the difference between those two flowers, one has a style or stigma that’s exerted and other inserted. Turns out this is a major controlling outcrossing and mating system as species evolve but also within species for controlling mating. There’s one major QTL for this in tomato, which we’ve isolated, which is a regulation in the control of a gene controlling.

And the third one, which we’re currently working on, is one that controls seed size. And we don’t often think of tomatoes as something domesticated for seed size, but as in many plants domesticated for monoculture, seed size increased dramatically when tomatoes are domesticated. And at first we thought it was because the fruit got bigger and the seeds got larger, but as we’ve mapped out the QTLs, those seem to be independent processes. There are several QTLs per seed size. The one that we have isolated appears to be a change in the regulation of a gene controlling

endosperm development, but the type of regulation is very surprising. It's a regulation caused by an insertion of a regulatory element in one of the introns, which we thought to be inert parts of the genomes, which changes the regulation of that particular gene.

So overall the picture is that most of these were changes in the regulation or the one is a change in the protein that causes regulation. So all of this comes down very closely on the side of... That's perfect timing, because I'm just finished up here.

So let me just tie this in to say that not only do we have an untapped reservoir of diversity, that Dr. McCouch talked about, that we have a chance to utilize, but that variation found in that material is extremely unique. You cannot create it in the lab overnight. It accumulates through large periods of time in evolution, and it's the great repository that we have yet to tap. It's unlikely to be recreated by mutagenesis or for that matter by GMO approaches, because it's something that's so unique in the evolutionary process. So I think one of the goals that we can have as a whole society is to first learn how this variation, what it is, and second how to utilize it more efficiently. And that in itself lays a foundation for plant improvement well into the future.