

GA: Thank you very much, Phil. Our final speaker in this session is Barbara Schaal. Barbara is the Spencer T. Olin Professor of Biology at Washington University, and a research associate at The Missouri Botanical Garden. As a plant evolutionary biologist, Barbara has contributed significantly to our understanding of the conservation implications of loss of genetic variation in isolated populations and also genetic issues surrounding invasive species. Her work has been singled out for distinctions in numerous ways. Dr. Schaal is a member of The National Academy of Sciences. It s my pleasure to introduce Dr. Barbara Schaal.
(Applause)

BIODIVERSITY, GENETIC DIVERSITY,
AND CONSERVATION

BARBARA SCHAAL, *Professor, Evolutionary
and Population Biology,
Washington University*

BS: Thank you very much. And thank you to the previous speakers, because they ve certainly set the groundwork for what I would like to talk about.

What I specifically want to talk about now is the use of genomics in plant conservation. And I want to address: What kinds of questions can we get at by using genomic markers? And I m going to be using a specific subset of markers and specific questions. The range of what one can do with genomics, of course, is far broader than what I could possibly cover.

And in the course of this I would like to use three specific examples. Conservation genetics is really an applied science, and so I want to use examples where we can really make some sort of decision about how to deal with conservation issues. And the first example will be looking at the wild genetic structure of a domesticated crop, and looking at land races of those crops. Another kind of diversity is, of course, agro-biodiversity, and we heard quite a bit about that yesterday. I would like to talk about a threatened prairie-plant species, and how we manage species that are threatened. And then, finally, something that s extremely important in terms of biodiversity, and that is invasive species. So we ve got two....

So when we talk about genetic variability, there s a number of very important reasons for looking at genetic variability and we will hear so much more about that. The first is that it is important to understand levels of variability, because it gives us an index of the current viability of species. Things such as the Florida panther have difficulties because they re inbred, and supposedly because they lost a lot of genetic variability, or they fixed detrimental variants. In addition, we need genetic variability, because it maintains the future evolutionary flexibility of a species. And these are the kind of standard reasons why people study genetic variability in plants ... and animals.

Now, genomic markers have really aided tremendously our understanding of the conservation genetics of both animals and plant species, and they provide an index of overall genetic variability. So we can use these markers something like allozymes, initially, in the case of cheetah, or RFLP s, or microsatellites ... or all of these host of different kinds of genomic markers to get a general index of genetic variability.

Likewise, we use this genetic index of variability to really implicitly get a handle on the amount of phenotypic variability. We assume that, in fact, genetic diversity equals biodiversity, very often, of phenotypes. And the assumption is

that the more genetic diversity that you have, measured by these neutral markers, the more phenotypic diversity that you'll have. And we'll look at that assumption.

And then, finally, most importantly, for a lot of us here, we can use these neutral markers to infer past evolutionary events. And two of the examples that I'm going to talk about, we're going to use neutral markers to try to reconstruct what's happened to species that are in difficulty.

So genomic markers in plants. The suite of markers that we use in plants are slightly different than those that are used in animal populations, and that has to do with differences in the nature of the genome in the structure and the organization of the genome. A whole set of very elegant studies in animals has been done on mitochondrial DNA variation, and in plants we don't have a mitochondrial DNA genome that varies in the same kind of manner than animals do. And so that's a tool that we don't have available to us.

Available to plant biologists are microsatellites; direct DNA sequencing which has become much more common in the last couple of years; and then some markers that survey wide levels of the genome: RAPDs random amplified polymorphic DNA and AFLPs amplified fragment-like polymorphic DNA. These are tools that basically use a PCR strategy to amplify up certain parts of the genome, and then to run, essentially, a series of fragments out on a gel, and the presence and absence of these fragments are scored.

They are markers that have been used extensively among plants, and not very much in animals at all and the reason is simply because of the differences in the genome. Animals have mitochondrial DNA, and plant biologists need to use these other markers.

So the first case study that I would like to talk about is looking at the question of: Does genetic diversity from neutral markers equal biodiversity? And the case study is in the plant *Manihot esculenta*. And we heard about *Manihot* yesterday. It's the tropical crop cassava. It's extremely important in the nutrition of the developing world. It is the most important crop in sub-Saharan Africa, and it's eaten primarily in tropical regions by the poorest of the poor. It sustains something like 600 million of the world's poorest people. And it has relatively as we heard yesterday, relatively few studies have been done on cassava, and it's considered an orphan crop although that picture is fortunately changing.

We were interested in determining what the wild progenitor of cassava was. When we started this work, way back in 1989, there were several hypotheses about what the wild progenitor species was for this extremely important crop but, yet, there had been no definitive work. We did a series of genomic studies molecular phylogeny reconstruction and we were able to pinpoint that it was domesticated from one of several possible species in Brazil. So we had a few candidate species, but we were not able to really pinpoint directly that.

So we then began to use a whole series of genomic markers to look at cassava and let me first show you a little bit about the crop. It looks like potatoes. These are tubers. It's grown for these underground tubers. It's a marvelous plant. It's resistant to drought. It grows on very poor tropical soils. It's something that requires almost no care. The farmers simply take a stem, and plant the stem in the soil, and six months later nine months later you have a very high harvest of these tubers.

This is a field of cassava in Africa. And in Africa, and in much of Asia and, also, most of South America it's grown in fairly large patches. And it's grown primarily as a source of starch.

We employed a series of genomic techniques, to try to determine where cassava was domesticated. Our molecular phylogeny reconstruction allowed us

to pinpoint that probably a wild species *Manihot flabellifolia* was a progenitor. Because there was close correspondence based on a whole series of different genomic markers about six or seven different DNA sequences between the domesticated crop and this wild progenitor.

But we then wanted to ask the question: Where, within the range of the wild progenitor, was domestication? Where did that occur? And this is the work of Ken Olsen, who was a graduate student at Washington University and is now at North Carolina State. And so what Ken Olsen did was to employ a series of genomic markers, looking at the wild progenitor species looking at as many populations as he could possibly collect and also looking at the world core collection of cassava. This is a representative sample of the biodiversity of cassava supposedly.

And what he found was really very interesting. The populations that are in red are all wild populations of this *Manihot* species that contain alleles that are found in domesticated cassava. The white square boxes are the domesticated progenitor populations, but that do not contain alleles found in the domesticated species. And then, finally, the circle is another group of plants another species of plants that's thought to hybridize, potentially.

And what you can see is that, very clearly, domestication appears at least in terms of the allelic diversity it appears to occur in one part of the species range. And these data are based on three separate nuclear genes so gene genealogy is from three separate nuclear genes as well as a suite of microsatellites. And all four data sets were completely consistent. So we feel, with some confidence, that, in fact, this use of genomics, in this wild progenitor, has pointed to where domestication occurred.

And domestication has occurred in the area that's on the interface between the Amazon forest and the *cerrado*. This is an area of very, very high biodiversity not only in plants but also in animals. And the *cerrado* is an area that hasn't received as much attention as the Amazon rain forest, but is very close in importance in terms of biodiversity. And the cassava, the wild cassava plants, live in gallery forest where the Amazon forest comes up along riverbanks, into the dry *cerrado* region.

So we were interested, then, in asking: Does, in fact, when we look at this area in Brazil where domestication occurred, is there any indication of high phenotypic diversity? And in this case we weren't asking about phenotypic diversity in the wild progenitor but, rather, we wanted to go to one further step. We wanted to look at phenotypic diversity in land races of this crop.

The rationale was and this is an old hypothesis that certainly doesn't hold was that at the location of domestication, probably the association between humans and this crop has been the longest. And, perhaps, in those regions where there has been this long association between humans and this early stage of domestication, perhaps the crop is used in different ways.

And one of the things that was interesting using a RAPD analysis, we were able to assess overall levels of genetic diversity. And what you can see is that the average diversity for crop varieties in southern Brazil which is the monoculture, the modern kind of agriculture the biodiversity was .006, based on conversion of RAPDs to nucleotides. And if you do the same measure if you look in the Amazon region the diversity is much higher. So this suggested that, in fact, both based on the location of domestication, and also based on the diversity of simply land races of cassava, perhaps there were interesting uses of cassava in the Amazon region.

My collaborator, Luiz Carvalho, from Cenargen in Brazil, then took some really very arduous field trips to the Amazon region and began to survey for the

use of cassava in these regions. The idea being that perhaps there were types of cassava, and uses of cassava, that were really unknown.

And what he found was really quite astounding. In the Amazon region, this area of domestication, there are many quote primitive land races, early land races, that are still used by the indigenous people, and they vary tremendously. They vary in pigment. And, in this case, you can see....

This is a cross section of the tuber. This is the standard starchy tuber that s used in modern agriculture. And there are varieties that are kind of purple; there s varieties that this is actually kind of pink; pigment sometimes is around the margin; and, in this case, there s pigment all through the cassava tuber. And we are beginning to analyze the biochemistry of these pigments. This is beta-carotene which, of course, is very exciting for nutrition in the developing world. This is lycopene, and there s lutein, and there s a series of different kinds of pigments that are incorporated into these land races.

Likewise, you can look at carbohydrate. And, again, this is the agrobiodiversity this earliest stage of domestication from this wild progenitor. This is simply a starchy tuber. And what Luiz has done is taken a Q-Tip with some iodine, and swashed it across. And you can see there s lots of starch here. The same has been done here, here, here, here and here and what you can see is, there s tremendous differences in pigment, and in starch deposition. And, in fact, again, carbohydrate analysis has shown that some of these varieties that are used in this area of domestication contain sugar rather than starch; contain phytoglycogen rather than starch; or contain novel kinds of biochemistry and carbohydrates.

So the point from this study is that the kind of genomic information that we have, when applied to plant species particularly some of the wild relatives of these domesticated crops can provide extraordinarily useful information, and pinpoint our collection activities, and allow us to, in fact, sample a tremendous amount of phenotypic diversity, biodiversity, simply by following what these neutral genetic markers show.

Now, the next example that I would like to use is looking at markers for conservation management. One of the most sad things in the United States is and, of course, one of the glorious things about the United States is that we ve had these beautiful prairies, and deep prairie soils. And the natural communities the prairie plant and animal communities have, in many cases, either gone extinct or are really threatened, and are just in marginal habitats. And just very small relics.

So what we are doing in this case is, we re looking at the genetics of a very interesting plant *Asclepias meadii*. It s sort of the poster child for prairie plants. It is a prairie plant it s extremely long-lived. Many prairie species are very long-lived plants 40, 50, 100 years. It is self-incompatible. That means it cannot set seed by putting its pollen on its own stigma. It can also spread vegetatively, and so it can reproduce by simply sending out stolon, sending out stems, and producing new stems. And it has experienced extreme habitat fragmentation. Its discoverer, Samuel Mead in 1972, made a comment about he was having trouble finding it. And this was back, I m sorry, in 1872.

This is what the plant looks like. It s a typical milkweed plant, and it has a typical kind of milkweed seed dispersal. This is the habitat. This is where Mead s milkweed initially was. All in this area very deep prairie soils. The triangles are known populations that have gone extinct since people began studying this, and the circles are populations that are still extant. And you can see there s a whole series of populations here. This is in the Flint Hills, and these are some rocky glade outcrops. And there s a few plants in deep prairie soils of Iowa, and

some in Illinois. This is actually an introduced population. So it's a species that is federally threatened is threatened.

Interesting. Most Mead's in fact, all Mead's milkweed populations are under management. They occur in prairie sites that are under conservation management. There are two main categories. The first are populations that are maintained by mowing. These are old hay fields. And even back in the late 1800s, many farmers recognized the beauty of the natural prairies, and they conserved particular areas, and used these fields specifically for hay.

This provides an interesting management for these populations, because when mowing occurs, it occurs in June. And at that point what happens is, that the farmer comes along and cuts off the flowering head, or the seed head, of the plant. And what that means is, that for many generations several generations there has been no or little sexual reproduction in these populations. Because they're removed the seeds are removed by the farmer.

The alternative is a much more natural kind of management, and that is burning. We think that vast amounts of the prairie are maintained by burning. And, so, in some of the conservation zones where Mead's milkweed grows, the populations are burned. And, in that case, burning is done in November or January, and it doesn't affect the life history of the plants.

If you look at these populations, the mown populations are very vigorous-appearing. They have many, many stems. But if you look at the amount of flowering if you take the plants out and you grow them in a common garden, and you look at the intensity of flowering the flowering is very weak in these populations, and it's very vigorous in the burned populations. But, just by appearance alone, the mown populations look quite lovely. They have more stems per unit area, etc.

So this shows you a typical burn population. This is the Ozarks of Missouri. This is to make Rob DeSalle homesick. And this is a glade population, so this is an area that's very rocky. It's much more reminiscent of the Southwest of the United States. And that's where Mead's milkweed occurs.

We used as genetic markers in this case random amplified polymorphic DNA. And it turns out that while this is a technique that has limitations, it's perfect for these kinds of studies. Because, basically, what it does is, it gives you a genetic fingerprint of each individual. And what we were interested in was looking at how many genetic individuals were in a population and, moreover, how big were big particular clones? Remember, we had individuals that can reproduce asexually.

So, in this case, you see these are two lanes that are identical? These were two separate stems that were collected, so they were separate plants, and we wanted to assay whether or not they were the same genetic individual. And you can see that there's a complete correspondence. Whereas, if you go over to this one, you can see there's a different fragment here, and here's a different. So we knew that this was a different genotype than this one. And that was the only thing that we were interested in for these data.

The bottom line here is that you can have populations that had many, many plants and let me read you an example. Fifty-three plants, a reasonable number of variable markers, but only three clones. So some of these populations can have a very large standing crop of plants, but very few genetic individuals. And in other cases you can have high numbers of genetic individuals so 21 in this population in Kansas.

And, basically, what this correlates to is that populations that are mown have very low genetic diversity, even though they have high individuals and populations that are burned have high genetic diversity. And this shows you a schematic of a typical situation in a mown population, where sexual reproduction

does not occur. This is a circle, a diagram, around each one of these, which is an individual stem, and all of these are the same clone. Likewise, here, all of these guys are the same clone. These are the same genotype, and here. And we find populations will have two, three sometimes five or six genotypes. But very few genotypes in the mown populations.

In contrast, if you go to a population that's maintained by burning, you get a very different genetic picture. And you can see lots and lots of different clones many, many different clones and genotypes within a population. And we can summarize that. If you look at the mown populations, 3.5 is the average number of clones. Whereas, the more natural management of burning, the average is about 11.7 clones. So there's a different genetic structure, which is really undetected till we did this.

Now, this has long-term and immediate consequences for the population. The long-term consequences are the evolutionary flexibility. But in a plant like Mead's milkweed, which is self-incompatible, the consequences of losing genotypes of having low diversity are immediate. This plant must outcross. There must be movement between one plant, one genotype, and another because the plant has a self-incompatibility system. So the only way that you can maintain sexual reproduction in a population, and get new genotypes, is to have enough diversity at the self-incompatible locus to allow seed set to occur.

And, in fact, in many of the populations that look very vigorous, that are maintained by mowing, there is no sexual reproduction. You can manually cross the plants, and reproduction does not occur. The only way you can get these plants to reproduce is by taking pollen from one population and introducing it into another. So the loss of genetic variability is extremely significant for these populations. So the management of these *Asclepias* populations turns out to have an overwhelming effect. And it is extremely important to maintain populations by burning, because they increase they maintain genotypic diversity, and they allow sexual reproduction to occur.

Now, one last thing, and I'll go through this very briefly. This is the use of genomics to look at an invasive species. This looks quite lovely, but this is a very nasty species. This is *Tamarix ramosissima* it's called salt cedar. And this is the work of John Gaskin, who's a graduate student at Washington University. The former work, on *Asclepias*, was done by Doug Hayworth and Marlon Bowles of The Morton Arboretum. This is John Gaskin's work.

And *Tamarix* is called salt cedar. It's an Old World genus. There's 54 species, and it's the second-worst invasive species in the United States. The worst is *Lythrum salicaria*, purple loosestrife. It has taken over at least a million riparian acres. It's a particular problem in the Southwest. It grows along waterways. It sticks down a very large taproot and starts pumping up water. And it extracts salts from the waters and puts the salt on the surface of the leaves. The leaves fall down you have a saline duff which completely eliminates biodiversity. So you have very, very low biodiversity; you also lower the water table.

It's really a plant from hell. And it's nearly impossible to eradicate. The only way you can really get rid of it You can't burn it; you can't put herbicides on it. The best way is to bulldoze it, and that's not very friendly. This shows you an example in Utah of what *Tamarix* does. This is just miles of salt cedar taking over this area. So a very, very serious problem for our biodiversity.

So the objective of John's work was to try to figure out: Where did this invasive thing come from? What was the source of the invasion? It turns out that we weren't sure what the species was in the United States, so there was a taxonomic question. Moreover, even if you thought you knew the taxonomy, USDA was trying to do it by control studies. And what they found was that they

would take *Tamarix* from one place in the U.S. and get an insect to eat it. But, then, if they went to another population and tried to get that same insect to eat the *Tamarix*, it didn't work. So there was some kind of heterogeneity, and so the idea was to try to find the source populations.

And we also wanted to determine if hybridization had anything to do with it. One of the horrible things that happens when we introduce species into the United States, into nonnative habitats, is, things that were once isolated from each other—species that were once isolated can come together and hybridize. And produce, in some cases, potentially invasive species.

So this shows you the invasions—a large invasion here. This is *Tamarix ramosissima*; this is *Tamarix canaryensis*, from the Canary Isles; and this is another species. This is the one that we're really worried about—this huge invasion. If you look—this is the source. This is where the wild species are. And so John has gone to Georgia, Kazakhstan, Iran, China—looking for wild *Tamarix*. And he's used, now, a different kind of an analysis. We've heard from a number of people about phylogeography basically takes a gene genealogy and lays it in a geographical context. And it's perfect—it's absolutely perfect—for answering these kinds of questions.

So we sampled the invasive populations in the U.S., to know what genotypes we had in the U.S.; sampled the *Tamarix* species from Eurasia that are candidates for the invasive; and then we used, in this case, phosphoenolpyruvate carboxylate—an intron, a nuclear gene—and we reconstructed the genealogy....

This is what the genealogy looks like. And, of course, you can't read it—you can probably not even see it. What each one of these steps is, is a single mutational step that leads to a haplotype. And when there's a box around the haplotype, that means that that is a haplotype that we observe. These are mutational events that we no longer can see. All of the red ones—which are located here, and here, and here, and here, and here, and here and here—are haplotypes that are found in the United States. Likewise, all of these dark ones—the black ones here—are things that are wild species from various parts of Eurasia.

And what's interesting is, if we look right here—this is a haplotype that's found in New Mexico, Colorado, Wyoming, California, Utah, Texas and Missouri. And it comes from China. It's *Tamarix chinensis*. There's another haplotype here which comes from Kazakhstan, but it's also found in California and Wyoming. And, in fact, in many places what we find—and this is from *Tamarix ramosissima*. This tends to be in the eastern part of Eurasia, this species, and this tends—I'm sorry. This is in the western part of Eurasia—this is in the eastern part of Eurasia.

And what we find in the United States is, not only do we have these two different things in the United States, but they come together and form a particular variety of *Tamarix* that contains genomes from both the species. So we have very strong evidence that hybridization occurs. So our data are consistent with a hypothesis of hybridization.

So what I hope that this shows is, simply, that we can use genomic studies not only to address sort of theoretical issues, but they can be extremely important in on-the-ground active conservation management. Both in animals, such as the Florida panther, and also in plant species.

Thank you.

(Applause)