

Plant, Microbial, and Insect Genetic Resources, Genomics,
and Genetic Improvement
FY 2001 National Program Annual Report

- Introduction
- Component I. Genetic Resource Management
- Component II. Genomic Characterization and Genetic Improvement
- Component III. Bioinformatics and Genome Databases

Introduction

National Program 301, Plant, Microbial, and Insect Genetic Resources, Genomics, and Genetic Improvement, is divided into three components: **Genetic Resource Management** (conserving a broad spectrum of genetic resources and facilitating their use in genetic improvement and scientific research); **Genomic Characterization and Genetic Improvement** (including molecular marker analyses, nucleotide sequencing, gene mapping, comparative and, to a limited extent, functional genomics, plant breeding, and genetic enhancement); and **Bioinformatics and Genome Databases** (developing new software tools for analyzing and managing genomic and genetic resource information and delivering the former via databases implemented on up-to-date computer networks). Together, these components are yielding breakthroughs in understanding genome composition, manipulating genetic material, and providing genetically diverse gene pools with sustained and enhanced agricultural value. During 2001, this program produced many important discoveries and advances. Some of these are described below, grouped by program component.

Selected Accomplishments by Component

Component I. Genetic Resource Management

The USDA/ARS collections of genetic resources represent the largest and highest quality source of raw genetic material for ensuring continued worldwide progress in genomic research and genetic improvement of agricultural commodities. One component of National Program 301 is the USDA/ARS plant, microbial, and insect genetic resource management program, which is conducted with a variety of university and corporate partners and with other national and international genetic resource programs, especially those at international agricultural research centers. There are very few publicly accessible germplasm collections maintained by the private sector, so all segments of worldwide agricultural research and genetic improvement rely on the USDA/ARS public collections of genetic resources.

USDA/ARS National Plant Germplasm System (NPGS). The NPGS is the premier plant genebank network in the world, managing more than 449,000 different samples invaluable for crop genetic improvement and scientific study. More than 119,000 of these samples (more than 27 percent of the total) were distributed in 2001 in response to more than 3,600 different requests. More than 35,000 different samples were increased in 2001 by controlled field cultivation. More than 80 percent of all the seed-propagated accessions are backed up in security storage at the USDA/ARS National Center for Genetic Resource Preservation

(NCGRP) at Ft. Collins, Colorado. In 2001, seven foreign and two domestic plant explorations were conducted successfully. The preceding activities enabled the scientific community to gain ready access to a wide diversity of genetic materials for research and crop improvement.

The website for the NPGS public database, the Germplasm Resources Information Network (GRIN), was queried thousands of times per day through the World Wide Web. An average of more than 500 new visitors queried GRIN every day in 2001, an increase in usage of approximately 10 percent. This high frequency of usage indicates that this database is a key resource for plant researchers and breeders worldwide.

USDA/ARS Microbial Germplasm Collections. The USDA/ARS microbial germplasm collections now manage more than 100,000 different strains of bacteria, fungi, and related microbes. These microbial collections serve as an important reference standard for microbial identification, as biocontrol agents, and as sources of natural products for pharmaceutical and other industrial uses. The largest collection at the ARS Microbial Genomics and Bioprocessing Research Unit, Peoria, Illinois, manages more than 80,000 different strains, including germplasm vouchers of patented microbes. The collection at Peoria recently enhanced its database, which is now accessed by more than 60 users per day.

Documentation for the superiority of liquid nitrogen storage for long-term plant germplasm preservation. The results of a 20-year-long comparison of lettuce seed viability under conventional -18C storage, as compared to cryopreservation in liquid nitrogen, indicated that cryopreserved seeds retain viability and vigor longer than those in conventional storage. USDA/ARS scientists at NCGRP, consider this the first documentation for the superiority of liquid nitrogen cryopreservation for long-term preservation of certain seed types.

Identification of wild sunflower populations with resistance to downy mildew. New sources of genetic resistance to downy mildew in sunflower are needed because sunflower varieties in France are now susceptible to new disease variants. Germplasm evaluation programs conducted by USDA/ARS scientists at the Red River Valley Agricultural Research Center (RRVARC), Fargo, North Dakota, identified wild sunflower populations with high levels of resistance, not only to downy mildew, but also to sunflower rust and mosaic virus. In the future, these materials will be crossed with highly-bred sunflower germplasm to develop the first sunflower hybrids resistant to all three diseases.

Disease-resistant hop germplasm. Production of hops (a lucrative specialty crop for beer production) in the United States is threatened by the devastating disease hop powdery mildew. USDA/ARS scientists at the ARS Horticultural Crops Research Unit, Corvallis, Oregon, identified from genebank collections, both male and female hop genotypes resistant to the strains of hop powdery mildew, currently found in Washington State. These materials will be incorporated into programs for breeding disease-resistant hops.

New method for *Petunia* germplasm identification. *Petunia* is one of the most important ornamental bedding plants in the United States with farmgate value in the tens of millions of

dollars. The continual demand for novel *Petunia* varieties places a premium on genetic diversity in this genus. USDA/ARS scientists at the Plant Science Institute, Beltsville, Maryland, used variants of a pigment synthesis gene as genetic markers to reliably distinguish ecotypes of several *Petunia* species, thereby more precisely estimating the genetic relationships among *Petunia* species.

New method for alfalfa germplasm identification. Optimal germplasm management and crop improvement efforts require efficient and effective means for measuring genetic variation. USDA/ARS scientists at the Grain Marketing and Production Research Center in Manhattan, Kansas, developed a new method for distinguishing alfalfa populations based on variability in chloroplast DNA. This new method has great potential for more rapid identification of different alfalfa genetic variants.

Development of GIS-assisted germplasm diversity assessments. USDA/ARS scientists at the ARS Plant Germplasm Introduction and Testing Research Unit (PGITRU), at Pullman and Prosser, Washington, demonstrated a significant association between patterns of ecogeographical diversity identified by Geographical Information Systems (GIS) and patterns of genetic variability in forage germplasm. This association suggests that GIS effectively identifies and characterizes ecogeographical regions as an aid to planning plant explorations that would capture a maximum amount of novel genetic variability.

Utility of core subsets tested with genetic data. In theory, an optimally-constituted core subset should encompass the maximum genetic diversity in a minimum number of accessions. USDA/ARS scientists at the Corn Insects and Crop Genetics Research Unit (CICGRU), Ames, Iowa, compared the isozyme allelic constitution of an initial sunflower core subset with the allelic content of a similarly-sized group of randomly-chosen accessions. The latter included alleles missing from the core subset, so the core subset was modified to include those variants. This study represents one of the best tests of the practical utility of core subsets, and its conclusions have affected how germplasm managers apply the former as a curatorial tool.

Evaluation of apple germplasm for disease resistance. USDA/ARS scientists at the Plant Genetics Resources Unit, Ithaca/Geneva, New York, collected genetically-diverse wild relatives of apple from China 5 years ago. Evaluations of these materials for resistance to apple scab, cedar apple rust, and fire blight, the three most serious diseases of apple, identified plants highly resistant to all three diseases. These materials are now being incorporated into apple genetic improvement programs.

Discovery of non-*Rhizobium/Bradyrhizobium* aquatic bacteria that are symbionts of grasses. Non-*Rhizobium/Bradyrhizobium* aquatic bacteria were found to form symbioses with legumes and with west African rice landraces. Building on these discoveries, further research may eventually elucidate the genetic systems in these bacteria that encode the ability to form symbioses, thereby enabling development of nitrogen-fixing symbionts for grain crops.

Component II. Genomic Characterization and Genetic Improvement

USDA/ARS researchers apply traditional and contemporary technologies to assess the inherent genetic potential of germplasm. To harness this genetic potential most effectively, novel, rapid, and more efficient methods for identifying useful properties of genes and genomes, for determining the function of genomic segments and for manipulating genetic material and genomic information, must be developed. Through genomic characterization, the structural and functional features of genomic segments or of entire genomes can be elucidated nearly simultaneously--providing knowledge that will help create new agricultural products, ensure food safety and security, and furnish new ways of boosting agricultural productivity. Consequently, sequencing genomes or genomic components of agriculturally important organisms and determining their function via structural genomic analyses are priorities for this component of National Program 301. Furthermore, the information generated by this National Program component, when combined with more effective breeding strategies, will help ARS scientists more rapidly attain high-priority genetic improvement objectives. All of the preceding tools and information furnish the bases for more effective genetic diversity assessments, diagnostic and varietal/strain identification tests, and superior methods of genetic improvement. Optimal progress in genetic improvement will result from the marriage of the preceding technologies with genetic resources, bioinformatics technology, and genomic/genetic information (see below). The ultimate result will be superior, genetically improved organisms for agriculture and additional genetic stocks to enable future research.

The ARS-CSREES cosponsored 1st International Comparative Insect Genomics Workshop was held in Washington, D.C., on October 29-30, 2001, with over 150 attendees participating. The achieved purpose was to form teams to develop insect genomics projects, and strategies to accomplish them.

New Strain of Honey Bees Resistant to Varroa Mites. America's domestic and feral honey bees are under attack from invasive mites. The ARS Honey Bee Breeding, Genetics, and Physiology Research Unit, Baton Rouge, Louisiana, laboratory has shown a sturdy new Russian honey bee to be somewhat resistant to these mites. The new bee is being released to bee breeders through a cooperative agreement.

Pierce's disease agent sequenced. Through an ARS cooperative agreement with Brazil, the American Vineyard Foundation, and the California Department of Food and Agriculture, a Brazilian team (which had previously sequenced a related strain from citrus) has sequenced the *Xylella fastidiosa* bacterium that causes Pierce's disease in grapevines. Companion work with the Department of Energy's Joint Genomics Institute has resulted in the sequencing of related bacterial strains from almond and oleander. Researchers are hopeful that a comparative genomics approach will yield clues to the bacterium's transmission and pathology with the goal of providing a solution to the organism's threat to California agriculture (\$33 billion to the wine industry alone).

Maize genome characterization. Precise knowledge of the location, composition, and

function of plant genes is key for accelerating the rate whereby crop productivity can be enhanced. For example, genetic differences in flowering time represent a major impediment to cultivating particular maize hybrids in different latitudinal zones without first conducting a lengthy breeding program focused on adaptation. USDA/ARS scientists at the Plant Science Research Unit (PSRU), Raleigh, North Carolina, have identified genes whose variants strongly determine the height of ears of maize and how rapidly after planting the maize flowers. These findings may help accelerate the progress of maize genetic improvement, and thereby increase the spectrum of genetic variability available for maize production.

Disease resistant broccoli. Downy mildew is one of the most destructive diseases of broccoli, the most valuable cole crop in the United States (\$500 million annually). USDA/ARS scientists at the Vegetable Research Unit, Charleston, South Carolina, produced an experimental line of broccoli with very high levels of resistance to downy mildew infestation and also identified easily-measured genetic markers closely linked to genes that confer resistance. The line may serve as a source of resistance genes for broccoli breeding, whereas the genetic markers may accelerate incorporation of the latter into varieties.

New soyfood soybean varieties. The market for soyfoods has expanded substantially in Japan and the United States, but there are very few U.S. soybean cultivars available that were bred specifically for this purpose. USDA/ARS researchers at PSRU, Raleigh, North Carolina, developed four new soyfood cultivars that will be grown in the United States for export to Japan. Such varieties may enable U.S. farmers to remain competitive in the export market for Asia.

New ornamental variety. Ornamental horticulture in the United States has a need for new cultivars that are unlikely to become invasive plants. USDA/ARS scientists at the Subtropical Horticultural Research Station, Miami, Florida, working with South African collaborators, released a new dwarf cultivar of African iris to the nursery trade in southern Florida. This cultivar's intrinsic properties would preclude its becoming an invasive species.

New superior hybrid citrus rootstock. Citrus is a billion-dollar industry in Florida, where about five million citrus trees die or become unproductive every year because of rootstock failure. USDA/ARS scientists at the U.S. Horticultural Research Laboratory, Ft. Pierce, Florida, released a new hybrid citrus rootstock that is longer-lived and more productive, thereby potentially saving Florida citrus producers millions of dollars each year.

Insect-resistant maize germplasm. Maysin is a plant compound that confers resistance to corn earworm (CEW), a pest that causes millions of dollars of lost profit annually in the southeastern United States. USDA/ARS scientists at the Crops Genetic and Breeding Research Unit, Tifton, Georgia, produced a maize germplasm line that contained ten times the normal amount of maysin and are distributing this material to research programs that seek to enhance resistance to CEW.

New genetic markers for accelerating cotton genetic improvement. Genetic markers can accelerate the progress of crop improvement, but the biggest limiting factor to their

application to cotton breeding is the limited number of publicly-available markers. USDA/ARS scientists at Starkville, Mississippi, identified more than 100 such markers that are actually parts of active genes and another two markers that are very closely associated with the important and difficult trait of resistance to nematodes. These new genetic markers may tangibly accelerate the pace of genetic improvement in a crop where yield gains have stagnated in recent years.

Superior cotton genetic map. Genetic maps are key tools for optimizing the progress of genetic improvement. USDA/ARS scientists at the Crops Genetic and Production Research Unit, Stoneville, Mississippi, statistically merged genetic maps (three developed at Stoneville) for each of the four U.S. Upland cotton genepools to form a more comprehensive consensus Upland cotton genome map. The genetic material for cotton is quite complicated, and superior maps may substantially assist cotton improvement efforts.

Superior new rice variety. Rice cultivars with higher yield, better grain quality, and superior resistance to diseases and insects are needed to increase and sustain profitability for U.S. rice producers. USDA/ARS scientists at the Rice Research Unit, Beaumont, Texas, released “Saber”, a new rice variety for the southern United States with unique, highly effective resistance to the serious diseases rice blast and sheath blight and also with superior yield potential and grain quality.

Potential resistance genes for barley and wheat scab. Fusarium head blight, or scab, is a devastating disease of wheat and barley in the upper midwestern United States. USDA/ARS scientists at RPVARC, Fargo, North Dakota, in collaboration with scientists at Kansas State University, isolated four genes from wheat that may confer at least partial resistance to scab. These genes were provided to other researchers for incorporation into highly-bred wheat lines via genetic engineering. Scientists at Fargo have inserted anti-scab toxin genes into barley that may confer scab resistance.

Rate of genetic contamination quantified for wheat. The cultivation of genetically-engineered wheat in the United States must be preceded by risk assessments, including the risk that the genetically-engineered genes spread beyond specific cultivated fields. USDA/ARS scientists at the Wheat, Sorghum, and Forage Research Unit, Lincoln, Nebraska, quantitatively documented the frequency whereby specific non-transgenic wheat types contaminated different nearby non-transgenic wheat types via pollen flow and seed admixture. This information will be extremely important for performing valid, scientifically-based risk assessments for introducing transgenic wheat into the United States.

New superior potato varieties. The farmgate value of potatoes is worth several billion dollars in the United States; but cost of production is high because of diseases, pests, and suboptimal quality. USDA/ARS researchers at the Small Grains and Potato Germplasm Research Unit, Aberdeen, Idaho, and university cooperators in the Pacific Northwest released several superior potato varieties that should increase producer profitability. Klamath Russet and Keystone Russet provide superior U.S. #1 yield for fresh-pack operations, whereas the high-yielding Silverton Russet is suitable for both fresh market and processing.

New disease resistant pinto beans released. USDA/ARS researchers at Prosser, Washington; Mayaguez, Puerto Rico; and Beltsville, Maryland, and university cooperators from Colorado, Michigan, Nebraska, North Dakota, and Washington State released 14 pinto bean lines (of various market classes) resistant to common bacterial blight, a disease of increasing importance in the United States.

New superior pea varieties. Peas are important rotational crops in cereal-based cropping systems of the U.S. Pacific Northwest. But many diseases, insect pests, and environmental stresses reduce their seed yield and quality. USDA/ARS scientists in Pullman, Washington, released two new varieties of dry pea with resistance to many diseases and improved seed quality. These superior new varieties are expected to command a significant share of the pea production acreage in the Pacific Northwest and afford improved profitability to producers.

Elucidation of genomic regions important for wheat flour quality. Milling and baking qualities are important traits for breeding improved wheat. USDA/ARS scientists at the Western Wheat Quality Laboratory, Pullman, Washington, identified genomic regions of wheat associated with these quality traits. That association will enable those regions to serve as markers for more easily selecting club wheat with optimal quality traits.

New varieties of lesquerella. Oil from the new crop, lesquerella, contains unique hydroxyfatty acids. Lower cost of production is required to furnish this valuable industrial oil to consumers at an attractive price. USDA/ARS scientists in the Western Cotton Research Laboratory, Phoenix, Arizona, released new lesquerella varieties with superior agronomic traits that should lower production costs.

Superior genetic markers for selecting for high-yielding maize. Grain yield in maize is strongly determined by the number of kernels per ear, which is in turn strongly dependent on the number of kernel rows per ear. USDA/ARS scientists at the Western Regional Research Center (WRRC), Albany, California, isolated and cloned maize genes that regulate kernel row number. Armed with this information, maize breeders could more readily select ears with particular row numbers for optimal yield.

Nematode-resistant sugarbeets. Root-knot nematode is one of the most serious pests of sugarbeets, causing millions of dollars of damage annually. USDA/ARS scientists at the Crop Improvement and Protection Research Unit (CIPRU), Salinas, California, released the first sugarbeet line resistant to root knot nematode to breeders for incorporation into production material.

Potato late blight resistance mapped. Late blight, the most devastating disease of potato, was a proximate cause of the Irish potato famine in the 19th Century. It still significantly decreases yield and profitability in U.S. potato production. USDA/ARS researchers at the Vegetable and Crops Research Unit (VCRU), Madison, Wisconsin, and their university collaborators mapped the gene(s) governing resistance to this disease to a specific chromosome in a wild relative of potato. Mapping will facilitate eventual isolation and cloning of this gene, which would in turn enable its transfer and manipulation in cultivated potato.

Completion of largest rye genetic map. For information about rye genes and genetic sequences to assist rye crop improvement programs, their location on the rye genome should be known. USDA/ARS scientists at the Plant Genetics Research Unit (PGRU), Columbia, Missouri, constructed the largest (containing the most genes and linkage units) genetic map of the rye genome to date. This new map will facilitate the process of marker-assisted selection in this crop.

Important sorghum genes mapped. Limited genetic diversity of U.S. sorghum varieties makes sorghum vulnerable to disease. The chromosomal locations for genes controlling agronomically-important traits such as disease resistance, grain quality, and pollen fertility (important for developing sterile lines for hybrid production) in sorghum were identified by USDA/ARS scientists at the Southern Plains Agricultural Research Center, College Station, Texas. This information will provide breeders with the means to improve sorghum more efficiently, thereby reducing economic losses due to pathogens and environmental extremes.

Genetic markers for increasing the efficiency of sugarcane breeding. Perhaps 40 percent of the offspring of sugarcane breeding crosses are actually self-pollinated progeny of the female parent useless for further breeding. This high rate of self-pollination seriously hampers progress of sugarcane genetic improvement. USDA/ARS scientists at the Sugarcane Research Unit, New Orleans, Louisiana, developed new genetic markers that effectively distinguish self-pollinated from hybrid offspring, thereby enabling earlier and more effective culling of self-pollinated plants.

Genetic marker for pollen sterility in onions. Relatively few pollen-sterility and fertility-restoring genes are available for onion and garlic, two of the most economically-valuable vegetables in the United States. The paucity of such genes hampers efforts to expand the genetic diversity of these crops. USDA/ARS researchers at the VCRU, Madison, Wisconsin, identified genetic markers for a key pollen sterility gene that were widely adopted by commercial onion and garlic breeders to accelerate the rate of developing hybrids.

Genes for resistance to virulent lettuce diseases identified. Big vein, dieback, and verticillium wilt are three of the most serious diseases of lettuce, the United States' most important salad vegetable. USDA/ARS researchers at CIPRU, Salinas, California, identified lettuce germplasm accessions, and populations of related wild species, with effective resistance to these diseases. Research on verticillium-resistant lines is most advanced with highly-developed breeding lines ready for deployment as needed.

Genetic markers for resistance to soybean disease. Breeding for resistance to brown stem rot, one of the most serious U.S. soybean diseases, requires expensive and time-consuming field assessments. USDA/ARS researchers at CICGRU, Ames, Iowa, and collaborators at Iowa State University identified genetic markers for genes that govern resistance to this disease, which should enable more rapid and effective development of disease-resistant soybeans.

Component III. Bioinformatics and Genome Databases

USDA/ARS researchers and their public-sector cooperators are developing new software tools for rapidly and readily analyzing and processing the torrent of genome sequence and map data issuing from various large projects which are analyzing crop and model plant genomes. The processed data are managed in the USDA/ARS genome databases, bioinformatics focal points for the collection and dissemination of information relating to genomes, genetic improvement, genetic diversity, and their study. The databases are accessible through the World Wide Web and contain research contacts, the sequences of genes and expressed sequence-tags (ESTs), genetic maps, description of defined genes and DNA markers, genetic control and other genetic information about economically important traits, metabolic pathways and associated enzymes, links between separate model species and agricultural species databases, and links to relevant genetic resource databases and their extensive data on accession traits. The high-quality databases and software tools provided by this National Program component are optimizing access to genetic/genomic information for research, breeding, and genetic resource management.

New ARS networks for bioinformatics research and development and genome databases. With the USDA/ARS Center for Agricultural Bioinformatics at Cornell University, Ithaca, New York, serving as a hub, the USDA/ARS effort in bioinformatics and crop genome databases (conducted in partnership with university and private laboratories) expanded its overall capacity. USDA/ARS scientists and collaborators at the Cornell University Computer Theory Center tested a prototype data warehouse for managing millions of individual cross-sequence comparisons. In cooperation with Cold Spring Harbor Laboratory, USDA/ARS scientists selected a system based on Ensembl for ready comparative genomic mapping of grasses, gramene. In cooperation with the University of Minnesota, USDA/ARS scientists assessed the utility of software components that may be assembled into a high-volume, high-speed sequence analysis data pipeline. The efforts will enhance the USDA/ARS genome database and bioinformatic capabilities to effectively manage and deliver to users the masses of sequencing data issuing from the sequencing of model species genomes and of gene segments from crops termed expressed sequence tags (ESTs).

Maize data integration enhanced. Isolation and analyses of genes controlling important agronomic traits in maize would be aided by databases that integrate information about gene sequences, gene function, and chromosomal location. USDA/ARS scientists at PGRU, Columbia, Missouri, and their university collaborators integrated data in the database MaizeDB so that the preceding sources of information are linked, and so that MaizeDB is more interoperable with other genome databases (e.g., ZmDB, gramene, GRIN, Genbank).

Wheat genes identified by their sequence. Identifying individual genes by unique tags (also known as ESTs) is an essential prerequisite for further genome analyses. USDA/ARS scientists at WRRRC, Albany, California, produced more than 70,000 individual ESTs for wheat, barley and their close wild relatives. This information represents more than 85 percent of such ESTs that are publicly available worldwide, thereby furnishing a highly valuable tool for accelerating the progress of public and private sector cereal genetics and breeding.

Sorghum genomic information elucidates the structure of the sugar cane genome. The size and complexity of sugarcane's genome greatly complicate construction of genetic maps and placing new genes on those maps. USDA/ARS scientists at the U.S. Pacific Basin Agricultural Research Center, Hilo, Hawaii, and collaborators at the University of Georgia, Texas A&M University, and the Hawaiian Agricultural Research Corporation used more highly refined genetic maps of the related crop sorghum, whose genome is much smaller and simpler, as templates for constructing a superior genetic map for the two important sugarcane species. The arrangement of genes on the sorghum and sugarcane genomes was somewhat similar so that information derived from the sorghum genome may help accelerate genetic research and crop improvement in sugarcane via marker assisted selection.