USDA
Agricultural Research Service

National Program 301 –
Plant Genetic Resources, Genomics, and Genetic Improvement

Action Plan 2013-2017

USDA AGRICULTURAL RESEARCH SERVICE

National Program 301
PLANT GENETIC RESOURCES, GENOMICS, AND GENETIC IMPROVEMENT

Action Plan [2013-2017]

Goal: National Program (NP) 301, *Plant Genetic Resources, Genomics, and Genetic Improvement*, supports research that expands, maintains, enhances, and protects the United States genetic resource and information base, and increases knowledge of the structure and function of plant genes, genomes, and of plant biological and molecular processes. Through innovative research tools and approaches, this National Program manages, integrates, and delivers vast amounts of genetic, molecular, biological, and phenotypic information to a diverse global clientele. The ultimate goals for the preceding efforts are to improve the production efficiency, yield, sustainability, resilience, health, product quality, and value of U.S. crops.

The activities conducted under this National Program address the need to:
- Ensure the long-term safety and integrity of our agriculturally valuable genetic resource collections and associated information;
- Develop novel approaches for analyzing complex traits, identifying favorable genes underlying those traits and deploying them in breeding stocks;
- Exploit new technologies that enhance traditional methods of genetic improvement;
- Genetically improve a broad spectrum of major, specialty, and new crops;
- Increase our knowledge of the structure and function of crop and microbial genomes and of plant biological and molecular processes; and
- Acquire, analyze, and deliver genetic and information resources to the public.

Genetic resources are the foundation of humanity’s agricultural future. The Agricultural Research Service (ARS) genebanks contain sources of resistance, tolerance, and adaptation to biotic and abiotic stresses and new genes to improve the quantity, quality, and production efficiency of our food, feed, bioenergy, fiber, and ornamental crops. To ensure those genes are available for research and breeding, ARS must continue to strategically acquire and conserve germplasm that contains them; to develop new screening methods for identifying favorable traits; to ensure that germplasm is distributed where and when it is needed; and to safeguard these collections for future generations.

ARS has taken a leadership role in developing and curating crop genomic and phenotypic databases, which will expand under this Action Plan. New tools are needed to efficiently capture phenotypic data and extract useful information, via innovative data-mining strategies, from the ever-increasing flow of data into these databases. Methods to interconnect databases containing diverse types of information are needed to more efficiently and effectively identify important properties of genes and genomes and apply that knowledge to crop improvement. In addition, these interconnected databases will enable researchers to better associate specific genes with agriculturally important traits and build upon genetic advances in one crop to speed genetic gain in others.

For major, specialty, and new crops, ARS – in close cooperation with diverse public and private-sector collaborators – will improve and broaden the genetic base of U.S. crops to reduce genetic vulnerability. NP 301 scientists will devise and apply new technologies to develop superior new crop varieties and enhanced germplasm and will accelerate the deployment of high-value traits into breeding populations. New genetic sources of key crop traits will be identified and incorporated into crop breeding lines and gene pools. New breeding theories and strategies that exploit the vast amount of available genotypic data will be developed to effectively capture the intrinsic genetic potential in germplasm. To do so, innovative tools will be developed for dissecting the architecture of complex crop traits. These tools will be applied to mapping and determining the function of underlying genes so that variation in them and in linked genomic sequences may be exploited as more precise genetic markers. This knowledge will also enable molecular and genetic variability to be associated with specific phenotypes to enhance the effectiveness of genomic selection, thereby accelerating and strengthening progress in crop genetic improvement.

Plant breeding has a long history of improving crops through phenotypic or marker-assisted selection, usually in the absence of information regarding underlying genes or gene networks that control key traits, or control how plants interact with their environment. Understanding the key biochemical pathways and the underlying controls for plant growth, development, and metabolism has been elusive. However, recent advances in plant molecular biology, genetic engineering, and physiology could create opportunities for improvement of crops by more precisely and rapidly manipulating the existing variability in plant properties, or by creating new variability when warranted. Knowledge of how to precisely modify biological mechanisms will contribute not only to advances in fundamental plant biology, but will also help identify potential unintended consequences of that manipulation.

In addition to conducting research, ARS and its university cooperators will continue to mentor and train the next generation of crop breeders, geneticists, bioinformaticists, genome researchers, and genetic and genomic resource and information managers. Developing the next cohort of researchers is particularly important, considering the projected strong demands for scientific personnel.

The Plant Genetic Resources, Genomics, and Genetic Improvement National Program (NP 301) is composed of three research Components (Figure 1):

- Crop Genetic Improvement
- Crop Genetic and Genomic Resources and Information Management
- Crop Biological and Molecular Processes

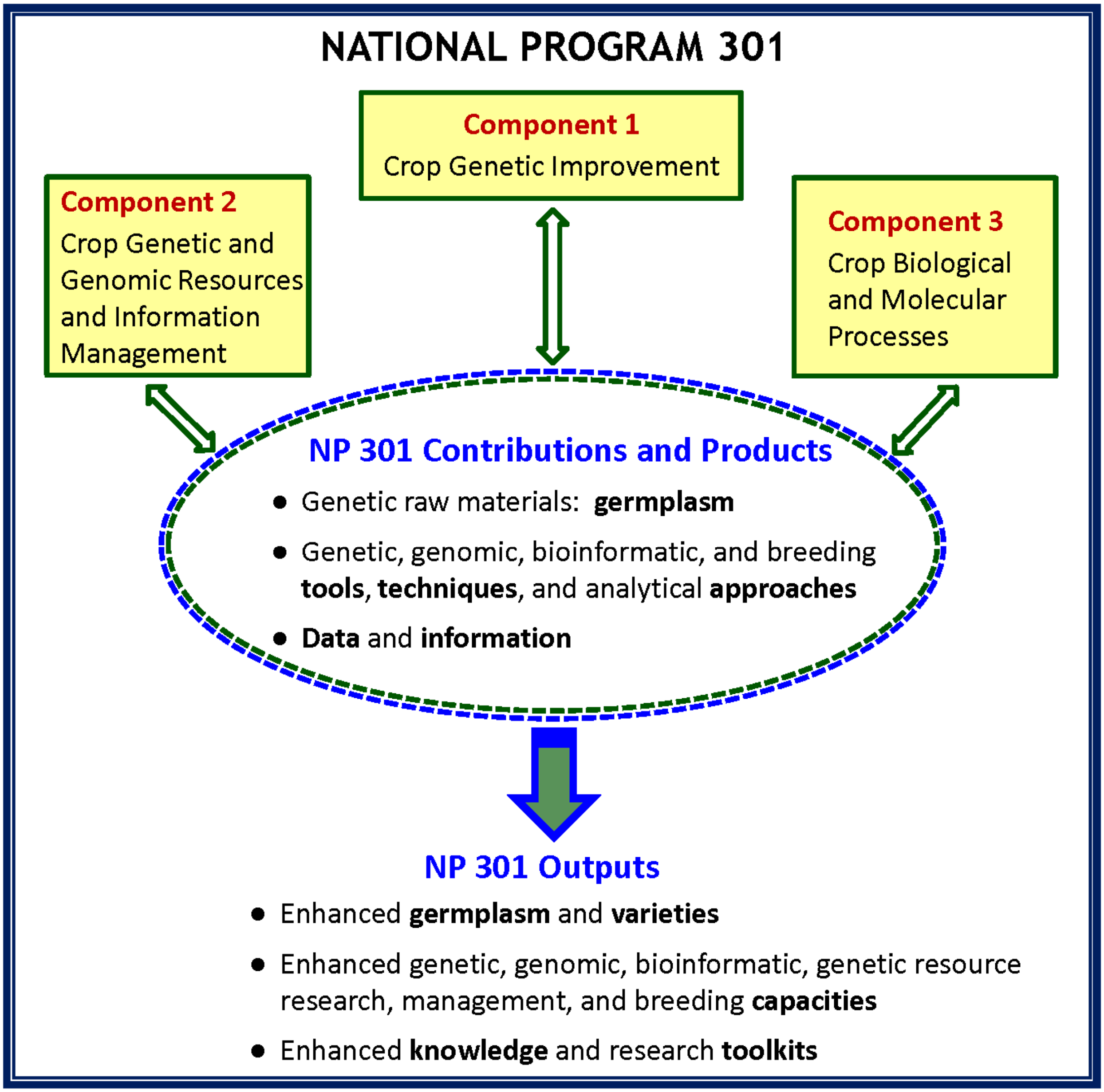

Figure 1: The three research components of NP 301 are interactive and collaborative and often blend resources to contribute towards accomplishments within specific National Program goals.

For organizational and administrative purposes, specific NP 301 research projects and resources are usually assigned membership in only one of these Components. Nevertheless, these categories of research components and their constituent Problem Statements do not act as barriers or “stovepipes” that impede research across the National Program. Rather, many of the individual research projects and resources contribute to the goals of multiple NP 301 Components, Problem Statements, or even several National Programs. The Anticipated Products for each of the six NP 301 Problem Statements are often critical for the success of research conducted under other Problem Statements.

Relationship of this National Program to the ARS Strategic Plan: Outputs of NP 301 research support the actionable strategies associated with the performance measure listed below from the *ARS Strategic Plan for 2006- 2011*, Strategic Goal 2, Enhance the Competitiveness and Sustainability of Rural and Farm Economies; Objective 2.2: *Increase The Efficiency Of Domestic Agricultural Production And Marketing Systems*.

> **Performance Measure 2.2.3:** *Expand, maintain, and protect our genetic resource base, increase our knowledge of genes, genomes, and biological processes, and provide economically and environmentally sound technologies that will improve the production efficiency, health, and value of the Nation's crops.*
>
> > **Target**: New technologies developed and used by ARS customers to increase production efficiency and enhance the economic value and quality of U.S. crop production while decreasing the environmental footprint of production systems.

Component 1: Crop Genetic Improvement

Essential to U.S. agriculture are plants that are higher yielding, durably disease and pest resistant, efficient in usage of all production inputs, and resilient to environmental extremes and changes in climate. Improved plant varieties are the products of innovative exploitation of genetic resources and efficient plant breeding strategies. Strategically coordinated research programs must expand the science of plant breeding, and develop improved germplasm to meet the challenges to the U.S. food, fiber, fuel, and feed supply. Effective genetic improvement of crops requires a broad genetic spectrum of "starting material" and knowledge of their intrinsic properties. Plant breeding approaches that integrate the most effective new methods and technologies with time-tested techniques will accelerate progress. A solid theoretical foundation of superior principles and methods underpins effective identification and evaluation of genetic variability for key traits, choice of breeding stock, and incorporation of traits into the latter. In addition, genetically and phenotypically characterized germplasm, experimental populations, and breeding stock are required to understand gene function and develop superior cultivars.

High-throughput genotyping and phenotyping methods are required for quantitative analyses of complex traits in diverse crop species. High-resolution genetic maps and full or partial genome sequences enable genetic markers to be identified and mapped. They enable the structure of complex traits to be elucidated to facilitate more effective genetic selection. Functional analysis of genes affecting such traits must be conducted. New methods of genomic selection and prediction enable genomic-assisted crop breeding approaches to deal with epistasis, polyploidy, differential heritabilities, genetic associations, and other underlying causes of such complexity. Application of genomic information from model plants and superior methods for identifying and introgressing exotic alleles into adapted genetic backgrounds will improve the efficiency of plant breeding.

Problem Statement: 1A: *Superior new crops, varieties, and enhanced germplasm.*

U. S. agriculture requires superior crops, varieties, and enhanced germplasm with higher yields; resistance or tolerance to biotic stresses and environmental extremes; improved agronomic or horticultural characteristics; and specific superior processing and product quality properties. These superior products must be adapted to a wide variety of multiple cropping systems and to competitive global market needs. Faced with a shrinking supply of land, labor, water, and inputs, U.S. crops must have enhanced water and input-use efficiencies and optimal production efficiency. New crops are required particularly to meet the needs for bioenergy and bio-based products, and current crops must be bred to meet new uses and market demands.

Research Needs:

Improved breeding and selection methods are needed to identify breeding lines and varieties that produce consistently high yields under a variety of climatic conditions and agricultural production systems. Improved methods for assessing the effectiveness of disease and pest resistance genes in diverse genetic backgrounds are needed to recognize durable resistance or tolerance to multiple diseases and pests in high-yielding plants. Superior approaches for introgressing resistance genes from exotic germplasm and progenitor species into adapted genepools will be developed. New annual and perennial crops with superior productivity under variable soil and atmospheric conditions must be bred. Crop plants with more desirable flavor, attractiveness, nutritive quality, and storage life should be developed. Varieties which are more efficient in their use of water and fertilizer are needed. These goals will be attained partially by improving above and below ground plant structures to maximize the genetic potentials for high yield and production efficiencies. High-energy, non-food plants capable of growth on underutilized agricultural lands will be bred as feedstocks for bioenergy production. Crop by-products will be evaluated for bioenergy applications or yield of higher value bio-based products. Value-added traits that improve end-use quality will be identified and incorporated into existing crops. Crops must be bred for new uses, including cultivation in protected environments, organic systems, small farm and dooryard production, and cultivation in cityscapes and non-traditional growing areas.

Anticipated Products:

- Higher yielding plants
- Plants with resistance or tolerance to diseases and pests
- Plants tolerant to environmental changes or extremes
- Plants with superior product quality
- Plants with enhanced water and input-use efficiencies
- Plants optimized for production efficiency
- Plants for bioenergy and bio-based products
- New uses for current crops

Potential Benefits:

The overall benefits of superior crop varieties and enhanced germplasm are a safe, plentiful, high quality food supply and a stronger, more secure agricultural industry. Crops with higher yields provide for food security, price stability, and a reduced need for food imports. Higher yields can help to stabilize rural economies, promote efficient use of arable lands, and increase farm-gate value. Genetic resistance can replace or supplement pesticide use and improve the competitiveness of U.S. agriculture in both traditional and organic farming sectors. Markets for reduced pesticide or pesticide-free agricultural products can be expanded. Crops with resistance or tolerance to multiple diseases and pests reduce pre- and post- harvest losses, and lower production costs. More durable resistance to diseases and pests and resilience to climate change can reduce the risk of catastrophic losses and improve food security. Crops with superior product quality can encourage consumers to eat healthy foods, thereby improving human health and creating increasing demand for high-quality U.S. crop products. Crops that utilize inputs more efficiently reduce production costs and enable cultivation in marginal environments. Improved plant architectures contribute to increased yield per unit area, thereby enhancing the economic viability of small producers and family farms, as well as large-scale commercial production agriculture. Productive, efficient, bioenergy crops adapted to marginal lands provide new opportunities for agricultural producers, ensure agricultural price stability, foster U.S. energy independence, and reduce environmental degradation associated with fossil fuel production and use. New uses for current crops will provide alternative markets for farmers and a variety of novel products for domestic and international consumers.

Problem Statement 1B: *Innovative approaches to crop genetic improvement and trait analysis.*

ARS conducts genetic and breeding research on a wide spectrum of crops, many of which require custom-tailored approaches. To meet the needs of rapidly expanding domestic and global populations, crop geneticists and breeders require not only the toolbox currently available from crop genetics and breeding, but also amalgamations of current knowledge with a suite of new basic and applied genomic tools and approaches. Methods, techniques, and knowledge gained from advances in genetic and genomic research – including improved, high-throughput phenotyping, new crossing schemes, and novel predictive computational tools – accelerate plant breeding and increase the specificity of breeding targets.

Building on these advances, ARS researchers exploit crop genetic resources (landraces, crop wild relatives, and older varieties) and elite breeding stocks as experimental populations for identifying and understanding the genetic bases for key traits. They elucidate the genetic diversity and population structure of crops (especially specialty crops) and crop wild relatives to reveal their potential for crop improvement. New research populations are also generated. These materials are characterized with improved and/or next-generation genotyping, phenotyping, and analytical approaches. The resulting new genetic tools and insights lead to testable hypotheses for novel plant breeding methods that maximize rates of genetic gain, and yield genetically enhanced populations improved for valuable agricultural traits, many of which are complex and quantitatively inherited. These traits are analyzed and catalogued with unprecedented levels of detail, and the range of effects and frequencies of relevant genes are functionally characterized, in both improved and germplasm populations.

A thorough understanding of genome structure and genetic variation enables more extensive trait analyses. Genetic, physical, and cytogenetic maps are integrated and, for polyploid crops, comparative analyses of syntenic genomes and shared genomic sequences improve the accurate assembly of physical maps. Improved genetic maps refine breeding techniques and approaches for genetic selection. Once advantageous alleles for specific traits are identified, genetic maps – highly saturated with genetic markers – facilitate marker-assisted selection for those traits.

Research Needs:

Both crop improvement and genetic/genomic research require knowledge of the genetic bases for the wide range of phenotypic variation maintained in breeding stocks, landraces**,** crop wild relatives**,** and other source populations. Understanding the genetic architecture of complex traits and the function of individual genes requires a diversity of experimental populations. Existing germplasm and breeding materials can serve as primary research materials, but in many cases they must be supplemented by new genetic stocks to enable functional genetic analysis. High-throughput genotyping, nucleotide sequencing, and efficient quantitative phenotyping technologies must be applied to analyze these populations.

High-resolution, sequenced-based genomic maps must be assembled to leverage knowledge of genetic regulatory pathways gained from model species and to identify genetic markers that enable populations to be enhanced for desired alleles. The function for these alleles must be understood so that their effects can be more efficiently measured and manipulated.

Innovative high-throughput, quantitative phenotyping methods are needed, because trait measurement now is often the critical bottleneck to genetic analysis and breeding progress. Phenotypic and genetic data from large populations must be combined in a coordinated effort to predict the performance of untested materials. As research capacity and accumulated knowledge permit, genomic selection and related breeding by prediction methods must be extended and evaluated empirically for more complex genetic architectures (e.g., many interacting loci, especially in polyploids), subpopulation structures, and a wide range of specialty and other crops which until now have lacked these genetic tools.

Anticipated Products:

- Genetically and phenotypically characterized germplasm designed for elucidating gene function and/or developing superior cultivars
- High-throughput genotyping methods and mapped genetic markers for priority traits in diverse crops
- High-resolution genetic maps and sequences for priority crops

- New phenotyping approaches for quantitative trait analysis
- Dissection of the genetic structure of complex traits and functional characterization of their constituent genes
- Methods of genomic selection and prediction
- Genomic-assisted breeding approaches for dealing with complex traits, and identifying and introgressing exotic alleles into adapted backgrounds
- Improved plant breeding techniques that apply knowledge from model plants

Potential Benefits:
The ultimate benefits from research conducted under this Problem Statement will be new tools, methods, and approaches for attaining crop breeding goals more rapidly and cost-effectively. More efficient and effective genotyping and phenotyping methods, superior functional genetic analyses, and the capacity to extend findings from model plants to crops will accelerate gene discovery and translate into cost-effective breeding tools. These technologies will also facilitate more comprehensive characterization of genetic stocks and other crop genetic resources, enabling their judicious incorporation into cultivar improvement and gene discovery programs. New specialized genetic stocks will not only furnish valuable research tools, but they can serve as a collaborative scientific focus for ARS and the broader crop genetic research community.

Products of gene discovery and other genomic/genetic research (e.g., genetic maps, markers associated with traits, and functionally characterized genes) will elucidate the genetic bases of and control for traits and accelerate their improvement. Improving the efficiency and accuracy of phenotypic evaluations will enable larger populations to be assayed, thereby improving the predictive power of genetic analysis. In particular, these research results might enable synergistic combinations of alleles to be selected, rather than selecting individual genes for their average effects. When feasible and applicable, genomic selection techniques and knowledge of gene interactions will be extended to specialty crops and those which are highly polyploid.

Component 1 Resources:
98 ARS projects in NP 301 address the research and infrastructure needs identified under Component 1. ARS lead scientists who are assigned to these projects include:

Maricopa, Arizona:	Terry Coffelt, Michael Gore
Stuttgart, Arkansas:	Anna McClung, David Gealy
Albany, California:	Olin Anderson
Davis, California:	Thomas Tai
Parlier, California:	Craig Ledbetter
Salinas, California:	Kelley Richardson, James McCreight
Fort Collins, Colorado:	Leonard Panella
Washington, D.C.:	Nancy Luria, Scott Aker
Canal Point, Florida:	Barry Glaz
Fort Pierce, Florida:	Kim Bowman
Miami, Florida:	Osman Gutierrez

Byron, Georgia:	Thomas Beckman
Dawson, Georgia:	Vacant
Tifton, Georgia:	Xinzhi Ni, C. Holbrook, Baozhu Guo
Aberdeen, Idaho:	John Bonman, Richard Novy
Kimberly, Idaho:	Carl Strausbaugh
Peoria, Illinois:	Patrick Dowd
Ames, Iowa:	Michael Blanco, Craig Abel, Reid Palmer, Marvin Scott, Randy Shoemaker
Manhattan, Kansas:	Robert Bowden
Houma, Louisiana:	Edward Richard Jr.
Beltsville, Maryland:	Mark Ehlenfeldt, Kathy Kamo, Margaret Pooler, Kathleen Haynes, John Stommel, Marcial Pastor-Corrales, Richard Olsen, Lisa Rowland, Perry Cregan
East Lansing, Michigan:	Mitchell McGrath, Karen Cinchy
St. Paul, Minnesota:	Vacant, David Garvin
Stoneville, Mississippi:	Brian Scheffler, Lawrence Young, Jeffery Ray, Jodi Scheffler
Mississippi State, Mississippi:	Johnie Jenkins, William Williams
Poplarville, Mississippi:	Timothy Rinehart
Columbia, Missouri:	Vacant, Michael McMullen
St. Louis, Missouri:	Ivan Baxter
Lincoln, Nebraska:	Scott Sattler, Robert Graybosch
Geneva, New York:	Gennaro Fazio, Peter Cousins, Christopher Owens
Ithaca, New York:	Edward Buckler, Jean-Luc Jannink
Raleigh, North Carolina:	Matthew Krakowsky, Thomas Carter, Gina Brown-Guedira, Jim Holland
Fargo, North Dakota:	Gerald Seiler, Thomas Gulya Jr., Prem Jauhar, Justin Faris, Lynn Dahleen
Wooster, Ohio:	Rouf Mian
Stillwater, Oklahoma:	Yinghua Huang, Kelly Chamberlin
Corvallis, Oregon:	Chad Finn, John Henning
Mayagüez, Puerto Rico:	Timothy Porch, Hugo Cuevas
Charleston, South Carolina:	Mark Farnham
Florence, South Carolina:	Philip Bauer
McMinnville, Tennessee:	Donna Fare
College Station, Texas:	Tommy Thompson, Robert Klein, Richard Percy
Lubbock, Texas:	Zhanguo Xin
Prosser, Washington:	Charles Brown, Duroy Navarre, Phillip Miklas
Pullman, Washington:	Camille Steber, Daniel Skinner, George Vandemark
Kearneysville, West Virginia:	Ralph Scorza
Madison, Wisconsin:	Philipp Simon, Shelley Jansky

Component 2: Crop Genetic and Genomic Resources and Information Management

The continued support and strategic expansion of crop genome databases, bioinformatic analytical capacities, *ex situ* genebank collections, and *in situ* dynamic genetic resource management programs are ARS priority goals. ARS is uniquely positioned to consistently maintain, enhance, and deliver genetic, genomic, and information resources to domestic and international research communities. This germplasm, associated information, and novel ways of transmitting, visualizing, analyzing, and manipulating genomic data constitute critical capacities, tools, and infrastructure support for essentially all phases of plant research and breeding.

ARS's actively curated, interconnected crop genome databases are keystones for long-term support of domestic and international crop breeding, genetic, genomic, bioinformatic, and biological research. Crop genomic information management capacities, tools, and infrastructure require enhanced, recurrent support, and improvement based on user surveys and working group recommendations. Furthermore, superior bioinformatic tools, particularly for managing high-throughput phenotypic and genotypic data, are continually devised for data analysis or data mining that extends knowledge generated from a particular crop to genetically related crops. These information resources effectively accelerate the pace of crop research and breeding when they are accompanied by user instruction and outreach programs focused on optimal database access and appreciation of the diversity of available data.

The highest priority for the ARS plant and microbial genebanks is to conserve genetic resources and associated knowledge in state-of-the-art genebanks and databases. These collections and descriptive information, vital to domestic and international national food security, are safeguarded in both primary and back-up sites. Genetic gaps in collection holdings are identified and priority genetic resources (especially for crop wild relatives and legacy materials) are acquired. Once in secure storage, their effective curation and use in research and breeding are encouraged by systematic characterization with genetic markers and by evaluation for priority agricultural traits to ensure that rare alleles vital for addressing critical threats are conserved and distributed. Without such knowledge, genetic resource management programs are handicapped and potential users are less likely to request the most relevant samples for research and crop improvement. More efficient and effective genetic resource and information management methods ensure long-term genetic integrity, health, and availability. Strong cooperative links and consultation with the National Plant Germplasm System's (NPGS) Crop Germplasm Committees (CGCs) foster high-quality genetic resource management plans that enable genetic resources and associated information to be delivered when and where needed. The origins and ecogeographical distributions for genetic resources transcend national boundaries; thus, effective conservation involves international partnerships and information exchange.

Problem Statement 2A: *Crop genomic information resources and bioinformatics.*

The recent development of extremely rapid and inexpensive DNA sequencing technology has transformed the study of plant biology. Research on major field crops, specialty crops, and model plants now generates billions of DNA base sequences from many thousands of plants from genetic stocks, mapping populations, genebank accessions, and wild populations. Collectively, these data are key to understanding and harnessing the genetic potential of crop genetic resources and breeding populations. ARS crop genome databases act as repositories for data generated by research projects funded by ARS, the USDA National Institute of Food and Agriculture (NIFA), the National Science Foundation (NSF), Department of Energy (DOE), and the National Institutes of Health (NIH), and also serve as research and training resources for students. In addition, new tools are under development for genomic, phenotypic, and genetic analyses that compile, analyze, interrelate, manage, and visualize the large datasets from diverse research projects.

These information resources serve domestic and international basic and applied researchers in government**,** universities, and industry with many types of research – genetic, genomic, molecular, physiological, and biochemical – and enable cross-referencing and querying information across disciplines and crops. At a time when crop research has become increasingly multidisciplinary, ARS plant genome databases effectively manage datasets and computational challenges that have increased exponentially in size and complexity.

Research Needs

During the last decade, technological advances that potentially can revolutionize crop research have led to a rapid increase in the size and complexity of molecular and phenotypic data sets. The corps of ARS crop genetic and genomic databases must be transformed into portals to deliver up-to-date bioinformatics tools, cyber-infrastructure, and knowledge linked genetically and at the molecular level to genome sequence information. New databases for mentoring and delivering genomic information for specialty crops are needed. A new generation of bioinformatic tools must be devised to accelerate the pace of research, plus safeguard and enhance genetic and genomic information. Ultimately, those data and tools must enable breeders to design superior crops rapidly and efficiently based increasingly on understanding how genotype and environment determines phenotype.

The management, analysis, and interpretation of data, rather than data production, are now the major bottlenecks in plant genomics research. Coordination and cooperation among researchers and data management experts are necessary to create software environments, which along with data discovery and exchange protocols, must enable interaction with diverse datasets. Researchers require centralized access points to the genomic, genotypic, and phenotypic data stored at various databases. Optimal exploitation of these resources, then, requires combining the information from many sources in innovative ways to discover and test fundamental plant biological principles. Researchers from multiple fields and disciplines must be able to find, understand, and effectively consult these resources in both traditional and novel ways.

Such innovative, synthetic approaches are not available off-the-shelf; instead, developing them presents a significant new challenge. Integrative computational analyses must be combined with high-throughput data generation methods to reveal fundamental biological principles and ultimately predicatively link genotype to phenotype. Developing and deploying standardized ontologies, nomenclatures, and data representations, as well as protocols for systems to interoperate, are prerequisites for associating genotype with phenotype. Crop genome databases must be expanded, curated, and implemented to enable divergent datasets to be accessible by a standardized set of protocols, including superior database interfaces and stand-alone informatics tools.

Methods are still lacking for exploring and interrelating large-scale genomics and phenotypic (trait) data sets with the environmental and experimental parameters needed to interpret those data. Multiple data repositories contain information pertinent to understanding the complicated, multigenic genetic control of important crop traits, but technologies for identifying and accessing specific data types from diverse databases are largely lacking. Thus, interoperable databases to manage large datasets, novel analytical methods, tools, and workflows are required. Increasingly large, diverse datasets must be transparently represented and exchanged among disparate research communities and data management operations.

Anticipated Products

- Actively curated, long-term, interconnected information resources that support crop breeding, genetic, genomic, and biological research.
- Bioinformatic tools for data analysis or data mining that can extend from a single crop to related species.
- New tools to manage high-throughput phenotypic and genotypic data.
- Expanded outreach to instruct users on database access and the diversity of available data.
- Database improvements resulting from user surveys and database working group recommendations.
- Expanded domestic and international research community utilization of bioinformatic tools and data.

Potential Benefits

The availability of integrated datasets that can be navigated and analyzed with ease will enable scientists to generate and associate new knowledge about phenotype, genotype, and genome interrelationships, and their interactions with environmental factors. Plant biological data will be available to a broader spectrum of scientists and stakeholders, thus democratizing data access and broadening participation in agricultural research. Such improvements will promote efficiencies in both time and cost that accelerate crop genetic research, breeding, and genetic resource management.

Problem Statement 2B: *Plant and microbial genetic resource and information management.*

Priority goals for ARS are the continued support and strategic expansion of *ex situ* plant and microbial genebank collections in secure, well-maintained modern facilities that maximize genetic integrity, health, and viability. Descriptive information about the genetic resources must be efficiently curated, stored, and made available via GRIN-Global, a Web-based advanced information management system.

When genetic gaps in collections are identified, new genetic resources are acquired through domestic and international field explorations and exchanges. The interdependence of nations for access to genetic resources necessitates international collaboration to share genetic resources and information, conduct research, combat genetic erosion, and strengthen collections, according to standards of access and benefit sharing established by both domestic legislation and relevant international agreements.

Distributing well-characterized, healthy germplasm of known genetic integrity to *bona fide* requestors maximizes the positive impact of the ARS's National Plant Germplasm System (NPGS) and its microbial genetic resources. Characterization and evaluation of genetic resources for priority traits generate data particularly important for effective acquisition and curation, and to enable requestors to identify the optimal germplasm for research, education, and breeding.

> ***Research Needs:***
> Best management practices must be developed, shared, and adopted to ensure the efficient, effective, long-term, and safe preservation of plant and microbial germplasm. Commonalities across crops must be determined to strategically develop methods that are broadly applicable for maintaining the genetic diversity during storage and regeneration, particularly during replenishment of aging materials in the base and active collections. Superior viability testing and monitoring protocols must be developed and implemented to increase the percentage and quality of stored accessions, and to lengthen the intervals between regeneration events. Methods for conservation and back-up for vegetatively propagated and non-orthodox seeded plants, such as *in vitro* and cryopreservation, must be improved and applied. New statistical genetic approaches, based on genetic relationships among accessions, are needed to optimize the efficiency and effectiveness of germplasm management. The conservation targets (genes vs. genotypes) must be considered while identifying priorities for efficiently managing and safeguarding different germplasm types (e.g., elite vs. wild).
>
> The GRIN-Global information management system must be implemented throughout the National Plant Germplasm System (NPGS) to provide genebank staff and germplasm users with an easy-to-use, browser-based interface to support genebank curation, and enable a wide clientele to select the optimal genetic resources for their needs through access to ample passport, phenotypic, and genotypic data. Healthy, true-to-type germplasm must be distributed when and where needed.
>
> Gaps in the genetic coverage of collections must be identified, especially for crop wild relatives and land races. To do so, taxonomic identity, geospatial distribution, genetic diversity and structure, systematic relationships, and life history information must be weighed. Pedigree and trait data will be especially important for some landrace and elite improved materials. Germplasm must be acquired through exchanges or plant

explorations, the latter focusing particularly on crop wild relatives. New uses for crops might increase the need for selected acquisition of genetic resources with specific traits. Ideally, genetic diversity assessments would be conducted either before or soon after acquisition to identify efficient management strategies and units for wild-collected materials and elite collections. Habitat, georeferencing, and sampling data documentation must be standardized to align with and promote database interoperability. GRIN-Taxonomy must be supplemented with additional information on crop wild relatives. As a complement to conservation *ex situ*, conservation *in situ* should be promoted when possible.

Prioritized descriptor and trait lists for phenotypic evaluation and genotypic characterization must be developed or enhanced to incorporate standardized ontologies and controls to facilitate cross-crop comparisons. Passport, environmental, geospatial, agricultural trait, and genetic data must be incorporated efficiently into GRIN-Global. The high costs for collecting phenotypic data require that agricultural traits are carefully prioritized according to users' needs. Different descriptors might be needed depending on the improvement status of accessions (wild vs. cultivar). New mechanisms for enabling integration of disparate data types should be assessed through pilot projects. New, less costly gene or sequence-based markers and high through-put phenotyping methods must be applied for collection characterization. The NPGS should coordinate accession genotyping with genomic sequencing and genetic marker analyses conducted under other NP 301 problem statements or by genomic research consortia and other genebank systems, to enable ready cross-dataset comparisons. Outcomes from genotypic and diversity assessments must be applied to collection management (e.g., to identify duplicate accessions). Core subsets and subsets for special purposes (e.g., reference, evaluation) should be defined.

Crop Germplasm Committees (CGCs), representing the community of plant germplasm users, provide key input to ensure the relevance of NPGS collections to research and breeding, and to thoroughly prioritize, plan, and effectively conduct genetic resource management objectives. Curators and CGCs must be fully informed of international access and benefit sharing regimes. NPGS practices should continue to facilitate *ex situ* and *in situ* conservation, international access to genetic resources and associated information, and appropriate benefit sharing resulting from their sustainable use. Financial and institutional resources can be leveraged through collaborations among countries and institutions to conserve, exchange, and sustainably use genetic resources and associated information.

Anticipated Products:

- Priority genetic resources and associated knowledge safeguarded in state-of-the art genebanks and databases.
- Genetic gaps in collections identified and priority genetic resources acquired.
- Genetic resources characterized systematically.
- Genetic resources evaluated for priority agricultural traits.
- More efficient and effective genetic resource management methods developed.

- High quality genetic resources and associated information delivered to requestors.
- International partnerships for genetic resources research and exchange expanded.
- Genetic resource management plans and Crop Germplasm Committees strengthened.

Potential Benefits:
High-quality, well-documented genetic resources with known genetic integrity and excellent vigor, viability, and health will be acquired, characterized, securely preserved, and promptly and widely distributed to support and benefit the scientific and educational stakeholder communities. Ultimately, these genetic resources will contribute key crop traits – or even new crops – to benefit farmers and consumers globally.

More specifically, germplasm accessions will be conserved under improved conditions for longer periods. Additional, high-quality accessions will be securely backed-up. Collections will be strategically expanded, becoming more comprehensive and better documented. Gaps in collections will be identified and filled by explorations and exchanges. Best management practices for genebank management will be implemented. Redundancy in global germplasm conservation programs will be identified and reduced. All of these achievements will promote genebank effectiveness, so that requestors receive precisely the germplasm they need.

Genetic resources can more readily address and solve global agricultural problems once information on accessions is more easily accessible through GRIN-Global. Germplasm will be regenerated and characterized for key traits critical to breeding and genomics programs. Genotypic and phenotypic diversity data will be applied to managerial decision-making, with the benefit of more efficient collection management. More effective transfer of detailed information to requestors will enable targeted selection of accessions, and stronger contributions to research and breeding objectives.

Genebank collections deliver benefits only when they are used. Genetic resource utilization will increase as relevant, timely, high-quality comprehensive characterization and agricultural trait data are generated and delivered to domestic and international requestors. International collaborations will improve the worldwide knowledge base of genetic resources.

Component 2 Resources:
37 ARS projects in NP 301 address the research and infrastructure needs identified under Component 2. ARS lead scientists who are assigned to these projects include:

Albany, California:	Olin Anderson
Davis, California:	John Preece
Riverside, California:	Richard Lee
Fort Collins, Colorado:	David Dierig, Christina Walters
Washington, D.C.:	Alan Whittemore
Miami, Florida:	Osman Gutierrez, Alan Meerow
Griffin, Georgia:	Gary Pederson

Hilo, Hawaii:	Francis Zee
Aberdeen, Idaho:	Harold Bockelman
Peoria, Illinois:	David Labeda
Urbana, Illinois:	Martin Sachs, Steven Clough
Ames, Iowa:	Randy Shoemaker, Candice Gardner, Carolyn Lawrence
Beltsville, Maryland:	Peter Van Berkum, Karen Williams, Edward Garvey, Dapeng Zhang, Gary Kinard, Amy Rossman, Richard Olsen
Geneva, New York:	Larry Robertson, Gan-Yuan Zhong
Ithaca, New York:	Doreen Ware
Corvallis, Oregon:	Kim Hummer
Mayagüez, Puerto Rico:	Ricardo Goenaga, Brian Irish
College Station, Texas:	Richard Percy, Larry Grauke
Prosser, Washington:	Stephanie Greene
Pullman, Washington:	Jinguo Hu
Madison, Wisconsin:	David Spooner
Sturgeon Bay, Wisconsin:	John Bamberg

Component 3: Crop Biological and Molecular Processes

ARS scientists conduct fundamental – often high-risk – research to generate an enhanced knowledge base and devise superior research tools that ultimately contribute to greater crop productivity and efficiency; better product quality and safety; improved protection against pests and diseases; enhanced tolerance to abiotic stress; and sustainable practices that maintain or enhance environmental quality. This Component emphasizes understanding the function of crop genes and devising the means for manipulating their expression. This research will elucidate the biological processes underlying crop productivity and quality, and develop new means for assessing risks associated with plant genetic engineering.

Historically, breeding has successfully improved crops through empirical approaches, often unaided by fundamental knowledge of plant biological mechanisms. The modern tools of biology offer new avenues for improving plants, either by exploiting existing variability in plant properties or by creating new variability when warranted. Knowledge of how plant biological mechanisms can be precisely modified will lead to new and potentially more effective approaches to crop breeding. As the scientific principles that link molecular and genetic phenomena to phenotype are understood, the pace of crop breeding should be accelerated and its efficiency enhanced.

Translating research at the molecular level into information that can solve agricultural problems demands an integrated approach that exploits knowledge of how variability in crop traits depends on gene expression and gene networks. This approach must provide the means to coordinate and integrate the fundamental with more applied methods to develop focused strategies for solving specific problems. Specific areas for emphasis include understanding how plants interact with the environment at the molecular, whole genome, and systems level; the control of plant growth and development at the genetic and epigenetic levels; and the control of plant metabolism and biochemical pathways.

Crop genetic engineering constitutes a major emphasis for ARS's fundamental plant research. It includes not only discovery research, but also efforts to improve biotechnology risk assessment methods and analyze potential unintended consequences of genetic engineering and transgenes in crop plants, on crop production, and on the environment. This involves developing superior new public domain genetic tools that improve the efficiency of conventional plant breeding and that also reduce the potential risks associated with transgenic crops.

Problem Statement 3A: *Fundamental knowledge of plant biological and molecular processes.*

Crop plants provide food, feed, fiber, and fuel as well as the products of their primary and secondary metabolism which are beneficial to health, nutrition, flavor, and industrial applications. To be a successful crop, the plant must efficiently convert resources such as light, water, and nutrients into high-quality and high-value plant organs and tissues. Crops must also function in terms of optimal patterns of architecture and development, biochemistry, and ability to respond to environmental stresses. The capacity to improve such traits is limited by gaps in our fundamental understanding of the underlying molecular and physiological mechanisms regulating plant metabolism, growth, development, and environmental responses.

Crop plants continually sense and actively respond metabolically to thousands of environmental cues. The effectiveness of these responses varies greatly throughout development, among varieties, and across geographical regions. How plants with similar DNA blueprints uniquely express their genes in response to different developmental or environmental cues is still poorly understood. Environmental extremes and more scarce water and other inputs require improved plant performance across diverse environments if the needs of consumers and producers are to be met.

Most important plant traits are often complex both at the genetic and physiological levels. Consequently, highly coordinated and integrated research into the regulation of plant biological processes is needed to improve crop adaption, yield, quality, carbon sequestration, and stress tolerance, while reducing costly inputs, expanding the area suitable for production, and protecting the environment.

Because of the complexity of plant responses to the environment, ARS research will apply expertise in genomics, proteomics, metabolomics, and epigenetics – including small RNA biology – to identify the basic mechanisms behind the genetic diversity in the preceding factors. The fundamental knowledge gained will identify the genes and associated mechanisms that will be harnessed for pyramiding of desired traits into our crops. Understanding the molecular mechanisms controlling key developmental transitions is needed to improve plant growth and architecture. ARS research will develop a better understanding of these processes and apply the knowledge gained to improve crop productivity, quality, product value, and sustainability. A highly coordinated integration of research is needed to develop knowledge of how metabolic and biochemical processes are regulated by internal and external factors, and how these processes can be modified to overcome factors that limit crop yield and crop quality.

Research Needs

Basic information on molecular, genetic, and physiological processes and their response to environment that underlie important crop traits is needed for more efficient and effective crop genetic improvement. Systems approaches drawing from next-generation sequencing, and proteomic and metabolomic techniques must provide large datasets for global gene and protein expression and metabolite levels both within and across crop species. In turn, these datasets will be mined to link the action of individual genes and gene networks to key plant phenotypes. Molecular, biochemical, and physiological phenotypes must be associated with genetic variation within and across crop and model plant species to enable the dissection and identification of the gene determinants underlying complex genetic traits that are priorities for crop breeding. Molecular, biochemical, and physiological approaches must be combined with the application of new algorithms and computational tools for data analysis and integration.

A more complete and detailed understanding of plant growth and developmental regulatory processes is necessary for developing plant breeding and biotechnology based strategies for maintaining and increasing crop productivity in response to a changing climate. Hormonal signals that trigger important developmental changes and the genes that control the signaling pathways must be identified. The potential for both identification of superior alleles and gene modification must be assessed for their role in altering the fundamental biological processes underlying improved crop productivity, superior quality, and enhanced resource use efficiency.

In addition, genes and gene networks underlying crop adaptation to biotic and abiotic stresses must be identified and the capacity for modifying them to enable crops to withstand stress or attack must be assessed. The molecular/biochemical basis for recognition and response between crops and both pathogenic and beneficial organisms must be determined and this information applied to identify important intervention points disrupting or altering interactions between crops and those organisms.

Integrating data from metabolite, gene expression, and protein profiling with new computational tools will be key for dissecting and elucidating metabolic pathways associated with the production of important phytochemicals. This should enable the identification of new sources of novel phytochemicals important for human and plant health and nutrition, and plant response to the environment. Hence, the identification of novel plant compounds must be related to the biosynthetic pathways, the genes that control those pathways, and the development of probes to enable rapid screening of plants for new sources of high-value compounds.

Anticipated Products

- Enhanced knowledge of how plants interact with the environment at the molecular, whole genome and systems level which is needed for developing resilient, well-adopted crops
- Determining how plant growth and development are controlled at the genetic and epigenetic levels, so that more productive crops can be bred
- Understanding the control of plant metabolism and biochemical pathways to enable higher yielding crops and new crop products to be developed

Potential Benefits

Knowledge of the molecular mechanisms underlying crop adaptation and plant-environment interactions will lead to higher crop productivity in unfavorable growth environments and improve the resiliency of agriculture to a changing climate. More effective breeding for pest and pathogen resistance will result from identifying and characterizing new sources of resistance integrated with a thorough understanding of plant function at the cellular and molecular levels. Knowledge of the mechanisms controlling both shoot and root architecture will improve crop productivity at all growth stages, including seedling vigor, nutrient acquisition, crop standability, and reproductive success. Understanding the molecular mechanisms controlling key developmental transitions will enable selection for optimal timing of seedling emergence, flowering, pollination, and seed maturation. Understanding the complex regulation of plant growth and metabolism will enable ARS scientists to improve crop plants for adaptation to changing climates, increase carbon sequestration, and value-added production. Knowledge of plant metabolism combined with understanding of plant growth and developmental processes can optimize biochemical pathways for synthesizing valuable chemical products.

Problem Statement 3B: *Plant biotechnology risk assessment.*

Genetic engineering offers tremendous promise for improving crop yields and crop protection; for increasing the efficiency and sustainability of production systems; and for providing the high-value and high-quality products required by the world’s consumers. These products can include foods with enhanced nutritional quality, plants with enhanced disease resistance, and new feedstocks for biofuel production. Except for a relatively few commodities, methods for efficient and effective crop genetic engineering have not been fully developed. Techniques for broadly applying these technologies to a wide range of crops are required.

Genetic engineering can produce novel gene combinations or arrangements in plant genomes, including recombining genes from distantly related species. Although these novel combinations offer great promise for crop improvement, the consequences of these manipulations require careful analysis to continue to protect human health and environmental safety. Research that integrates the development of efficient genetic engineering technologies, applicable to a wide range of crops and production systems, with robust and practical risk assessment strategies, will help advance the application of these technologies. In addition, this research will provide critical support to regulatory agencies charged with ensuring the safe utilization of genetically engineered crops, even as new genetic engineering technologies are developed. The intersection of genetic engineering, genomics, and plant breeding approaches will yield novel technologies that genetically modify a plant directly**,** but which generate improved plants not considered to be genetically engineered. This approach addresses issues with the coexistence of genetically engineered and non-genetically engineered crop production, and potentially creates a path to market for new products. These technologies are at early stages of development and require continued critical assessment and refinement.

Research Needs

Techniques from systems biology, next-generation DNA sequencing, and non-targeted metabolomic and proteomic profiling must be assessed and applied to expand the scale and depth, but reduce the cost, of crop genetic engineering and related research technologies. These technologies are also needed to understand how transgenic technologies change plants, their products, and their natural communities, as compared to effects from conventional crop improvement. Likewise, improved statistical and bioinformatic approaches for modeling gene, metabolite, protein, and community networks must be applied to analyze the potential impacts and interactions of crop breeding technologies, at both the micro- and macro- scales, on conventional, organic, and genetically engineered crops, coexistence of different production systems, and gene flow between populations.

Novel biotechnological approaches for the production of genetically engineered crops are required to surmount technical limitations, and to address the specific potential risks represented by genetically engineered products. Such approaches include:

- Methods and molecular tools for generating transgenic crops that enable precise control of transgene expression;
- Integration of DNA into defined sites in the genome;
- Selective removal of unwanted transgenic DNA from the genome post-transformation;
- “Pyramiding” multiple transgenes for transfer of multigenic traits; and
- Application of plant DNA elements for new genetically engineered product development.

Genetic engineering technologies have revolutionized the production of a few crops, but have not been applied to the vast majority of crops. New innovative approaches must be developed for genetic engineering technologies that integrate genetic transformation, genomics, and breeding, leading to improved varieties that are not categorized as genetically engineered, that can be rapidly deployed, and that address the coexistence of transgenic and non-transgenic crop production.

Anticipated Products

- Improved biotechnology risk assessment methods for the analysis of potential unintended consequences in crop plants, on crop production, and on the environment
- Improved public domain tools for producing transgenic crops

- Genetic engineering technologies that can serve as adjuvants to improve the efficiency of conventional crop breeding

Potential Benefits

This research will establish a basis for interpreting variability in global gene and protein expression in transgenic as compared to non-transgenic plants, as well as shifts in metabolism associated with transgenic and/or conventional crop breeding. Network analyses will estimate the positive, negative, and neutral consequences to individual plants, and to native and crop communities from genetic improvement. That knowledge, and improved public domain genetic tools and strategies for producing transgenic crops will address potential risks and provide critical information to regulatory agencies overseeing the safety of transgenic crops. These resources will enable broader application of this essential technology to improve both major commodities and specialty crops. The application of new genetic modification technologies to crop breeding, which enable genetically improved plants that are not categorized as genetically engineered, would enable broader application of genomic and biological knowledge to crop breeding. These innovative strategies would promote the integration of genetic engineering, genomic, and crop biological knowledge bases.

Component 3 Resources:

37 ARS projects in NP 301 address the research and infrastructure needs identified under Component 3. ARS lead scientists who are assigned to these projects include:

Albany, California:	William Belknap, Ann Blechl, William Hurkman, Frank Harmon, Sarah Hake, Sheila McCormick, Jennifer Fletcher
Gainesville, Florida:	Prem Chourey
Hilo, Hawaii:	Dennis Gonsalves
Aberdeen, Idaho:	Victor Raboy
Urbana, Illinois:	Donald Ort
Ames, Iowa:	Roger Wise
West Lafayette, Indiana:	Karen Hudson, Richard Shukle
New Orleans, Louisiana:	David Fang
Beltsville, Maryland:	Autar Mattoo, Benjamin Matthews, Anna Smigocki, Savithiry Natarajan
St. Paul, Minnesota:	Carroll Vance
University, Mississippi:	Stephen Duke
Columbia, Missouri:	Jan Miernyk
St. Louis, Missouri:	Yong-Qiang An
Ithaca, New York:	James Giovannoni, Leon Kochian, Owen Hoekenga
Raleigh, North Carolina:	Robert Upchurch
Fargo, North Dakota:	Karen Fugate
Lubbock, Texas:	John Burke
Kearneysville, West Virginia:	Michael Wisniewski
Madison, Wisconsin:	Cynthia Henson, Mitchell Wise, Ronald Skadsen, Johanne Brunet