National Program 301
Plant Genetic Resources, Genomics, and Genetic Improvement
Action Plan 2018-2022
Mission
To harness the genetic potential of plants to transform U.S. agriculture.

Vision
To be a global leader in research in plant genetic resources, genomics, and genetic improvement by providing knowledge, technologies, and products to deliver high crop yields and product quality, reduce global food insecurity, and decrease the vulnerability of global agriculture to destructive diseases, pests, and environmental extremes.

Introduction
Crop plants underpin U.S. agriculture and food security. Increased crop yield, product quality, and production efficiency are required to safeguard the Nation’s agricultural economy and food supply. This national program addresses the preceding critical needs for increased crop productivity by providing crop plants with higher inherent genetic potential. To do so requires continuous crop genetic improvement through more efficient and effective plant breeding, which exploits sources of new genes and traits from the Nation’s genebanks, leading-edge breeding methods, data-mining, bioinformatic tools, and incisive knowledge of crop molecular and biological processes. ARS breeders, plant scientists, computational biologists, programmers, and many others are engaged in a global quest to transform and accelerate the pace of innovation in plant breeding and crop production.

USDA Agricultural Research Service (ARS) places a high priority on realizing the products and accomplishments described in this National Program (NP) 301 Plant Genetic Resources, Genomics, and Genetic Improvement Action Plan 2018-2022, recognizing that without NP 301, “the risk and vulnerability to U.S. and world agriculture would be substantial.”1 NP 301 advances the ARS goals to protect, expand, and enhance the crop genetic resource base of the United States; increase scientific knowledge of crop genes, genomes, biological processes and systems; and deliver economically and environmentally sound technologies that improve the production efficiency, quality, health, and value of the Nation’s crops. At the core of its focus, ARS research in NP 301 will harness the genetic potential of plants and translate that potential into crop products that generate economic gain and consumer benefits. This National Program will contribute strongly to meeting the anticipated global demands for food, feed, fiber and

1 Quotation taken from the NP 301 Retrospective Review (June 27-28, 2016) Panel presentation
renewable energy and to meeting the USDA-ARS Grand Challenge to “Transform agriculture to deliver a 20% increase in quality food availability at 20% lower environmental impact by 2025.”

This National Program will develop, and effectively apply, new knowledge of crop genes, genomes, and the control and expression of genes, to accelerate the rate of genetic improvement and product quality improvement through traditional and novel plant breeding methods. Through innovative research tools and approaches, this National Program will manage, integrate, and deliver vast amounts of raw genetic materials (genetic resources); superior cultivars; and genetic, molecular, biological, and phenotypic information to a diverse global clientele. The ultimate goals for these efforts are to improve the production efficiency, yield, sustainability, resilience, healthfulness, product quality, and value of U.S. crops.

More specifically this National Program addresses the needs to:

- Ensure the long-term availability and integrity of the Nation’s plant and microbial genetic resource collections and associated information by safeguarding them in state-of-the-art genebanks;
- Identify and fill gaps in plant and microbial genebank collections, and acquire, characterize, and deliver high quality genetic and information resources to researchers, breeders, producers, and consumers;
- Devise new methods to accelerate discovery of new traits from the genetic variation in genebank collections, verified through novel, efficient phenotyping and genotyping approaches;
- Develop novel approaches for recombining genomes, conducting effective trait introgression, and applying production systems information to plant breeding;
- Produce higher yielding diverse crop plants that use water and other inputs more efficiently and contain genetic protection against crop production barriers, such as diseases, pests, and environmental extremes;
- Apply biotechnological and genetic engineering methods to a broader variety of crop species, and develop new methods to address their potential unintended effects on production systems;
- Increase knowledge of the control of plant growth and development, the influence of microbiomes on crop performance, and ways to enhance food quality and nutritional value at the genetic, molecular, and physiological levels;
- Increase knowledge of how crop plants interact with environmental factors at the molecular, whole genome, and systems levels;
- Maintain well-curated, interconnected information resources linked with databases that effectively safeguard and deliver huge volumes of genetic and trait information; and
- Construct high throughput phenotyping and genotyping capacities and efficient bioinformatics tools for data analysis and mining.
ARS’ National Plant Germplasm System, a major component of NP 301, contains 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. Conserved genetic resources, including crop wild relatives, play a critical role in crop genetic improvement and genomic research by contributing to the long-term safety and assurance that genetic diversity exists for future needs, especially considering the threats to genetic diversity, coupled with the uncertainty of future environmental extremes and changing climates. To ensure access to those genes for future research and breeding, ARS must continue to strategically acquire and conserve germplasm that can be exploited when necessary. ARS must also strive to develop efficient screening methods for identifying favorable traits, ensure that high quality germplasm is distributed where and when it is needed, and safeguard these collections for future generations.

In the plant breeding arena, the long-term goal of NP 301 research is to develop and release improved germplasm that includes breeding and pre-breeding materials, or superior cultivars for incorporation into commercial breeding programs or directly into commercial production as the crop species and specific commercial requirements might dictate. In many instances, trait-specific considerations will be important and will include high yields, disease and pest resistance, improved processing, and nutritional quality. NP 301 scientists will pay special attention to potential nutritional trade-offs as they consider selection for priority traits such as processing and shelf-life.

To attain higher crop genetic gain over the long-term, and to achieve other relatively short-term goals, breeding programs will incorporate innovative methods to modify genomes and devise new ways to incorporate exotic genes to develop higher yielding crop plants with desirable traits such as water and input use efficiency, and protection against diseases, pests, and environmental extremes. NP 301 scientists will devise and apply technologies such as biotechnology and genome editing to develop superior new cultivars and enhanced germplasm, and to accelerate the deployment of high-value traits into breeding populations.

Enabled by the NP 301 research that generates a fundamental understanding of the genetic and physiological bases for plant growth and development, plant breeders will devise more innovative methods to achieve sustainable yield increases to address environmental extremes and changes in climate, and to better understand the role of microbiomes in crop production systems in relation to phenotypic expression of priority traits. Also, knowledge of how various phenotypic responses are manifested at the biochemical and molecular levels will generate new insights that can be translated into superior traits for breeding programs. NP 301 research will develop innovative tools to foster innovative breeding strategies for designing crop plants with optimal architecture and efficient response to various environmental stimuli above and below ground.
As biotechnological advances continue, NP 301 research will deliver scientific assessments and technical leadership to support coexistence of classical and biotechnologically engineered crops. Evaluation of potential unintended consequences from the biotechnology approach, with regard to effects on crop plant biology, crop performance, and the environment, will be coupled with development of coexistence strategies. NP 301 scientists will devise ways to increase the preciseness of genetic engineering, such as precision genome editing.

NP 301 research will continue to lead the development and curation of crop genomic and phenotypic databases, and to devise ways to make the information more widely accessible to users. Those efforts will expand under this new Action Plan. NP 301 research will develop data management and exchange protocols to accommodate efficient access, analysis, and integration of these diverse sources. NP 301 scientists will develop bioinformatic solutions supporting research and breeding for the broad range of U.S. crops. NP 301 database curators will formulate efficient strategies to assure continued maintenance of the databases, with adequate infrastructure for storage, computation, and data movement. Standards will be developed for data collection and storage, and for data management processes, such as preservation and archiving.

New tools will be constructed to efficiently capture phenotypic data and to extract information, via innovative data-mining strategies, from the ever-increasing flow of data into these databases. New approaches for interconnecting databases containing diverse types of information will more efficiently and effectively identify important properties of genes and genomes and apply that knowledge to crop improvement. Furthermore, NP 301 scientists will develop better database interconnections that link the Germplasm Resource Information Network (GRIN)-Global with crop databases, and enable users to better associate specific genes with agriculturally important traits and to build on genetic advances in one crop to speed genetic gain in others.

In addition to conducting research, NP 301 scientists and 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 those scientists is particularly important, considering the projected strong demands for such personnel.

The Plant Genetic Resources, Genomics, and Genetic Improvement NP 301 comprises four research Components (Figure 1):

1. Crop Genetic Improvement
2. Plant and Microbial Genetic Resource and Information Management
3. Crop Biological and Molecular Processes
4. Information Resources and Tools for Crop Genetics, Genomics, and Genetic Improvement
The four research components interact to meet the overall goals of NP 301. In this interactive web, the Anticipated Products and Accomplishments for each of the NP 301 Components are often critical for the success of research conducted under other Components. However, for organizational and administrative purposes, each NP 301 research objective within a research project is usually assigned membership to only one of the four Components. Overall, many individual research projects and resources contribute to the goals of multiple NP 301 Components, and even other National Programs.

**Relationship of this National Program to the ARS Grand Challenge**
The NP 301 Action Plan will be a key contributor to the ARS Grand Challenge – to “Transform agriculture to deliver a 20% increase in quality food availability at 20% lower environmental impact by 2025.” Achieving this Grand Challenge will require coordination and collaboration among NP 301 scientists across Research Units, as well as with scientists in other National Programs. This coordination may involve common objectives in multiple NP 301 Project Plans that will serve to solve critical problems outlined in this Action Plan, and also to tackle the broader target goals of the ARS Grand Challenge.
**Relationship of this National Program to the USDA Strategic Plan**

This Action Plan will direct research that supports five objectives in the [USDA Strategic Plan for FY 2014-2018](https://www.usda.gov), under:

**Strategic Goal 1. Assist rural communities to create prosperity so they are self-sustaining, repopulating, and economically thriving.**
- Objective 1.3 – Contribute to the expansion of the bioeconomy by supporting development, production, and consumption of renewable energy and biobased products.

**Strategic Goal 2. Ensure our national forests and private working lands are conserved, restored, and made more resilient to climate change, while enhancing our water resources.**
- Objective 2.2 – Lead efforts to mitigate and adapt to climate change, drought, and extreme weather in agriculture and forestry.

**Strategic Goal 3. Help America promote agricultural production and biotechnology exports as America works to increase food security.**
- Objective 3.1 – Ensure U.S. agricultural resources contribute to enhanced global food security.
- Objective 3.2 – Enhance America’s ability to develop and trade agricultural products derived from new and emerging technologies.

**Strategic Goal 4. Ensure that all of America’s children have access to safe, nutritious, and balanced meals.**
- Objective 4.4 – Protect agricultural health by minimizing major diseases and pests to ensure access to safe, plentiful, and nutritious food.


This NP 301 Action Plan will direct research that supports:

Goal 1 – Sustainable Intensification of Agricultural Production: Subgoal 1A. Crop and Animal Production; Subgoal 1C. Crop and Animal Genetics, Genomics, Genetic Resources, and Biotechnology.

Goal 2 – Responding to Climate and Energy Needs: Subgoal 2A. Responding to Climate Variability; Subgoal 2B. Bioenergy/Biofuels and Biobased Products.
Relationship of this National Program to the ARS Strategic Plan for 2012-2017

This NP 301 Action Plan will direct research that supports:

Strategic Goal 3.1 – Protect, expand, and enhance the United States’ crop genetic resource base, increase scientific knowledge of crop genes, genomes, biological processes and systems, and deliver economically and environmentally sound technologies that improve the production efficiency, quality, health, and value of the Nation’s crops.

Performance Measure 1.3.1: Develop knowledge, strategies, systems, and technologies that maximize the production efficiency of our annual, perennial, greenhouse, and nursery cropping systems. Develop new technologies and tools contributing to improving these systems to meet current and future food crop production needs of diverse consumers, while ensuring economic and environmental sustainability and production efficiency, health, and value of our Nation’s crops.

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 of supplying the U.S. with food, fiber, fuel, feed, and ornamentals. Effective genetic improvement of crops requires a broad genetic spectrum of raw 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 stocks, 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 breeding exotic alleles into adapted genetic backgrounds will improve the efficiency of plant breeding.

**Problem Statement 1A: Trait discovery, analysis, and superior breeding methods.**

ARS conducts genetic and breeding research on a wide spectrum of crops, many of which require custom-tailored approaches. To meet the needs of rapid domestic and global population growth, 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 must include improved genotyping and phenotyping protocols. These protocols will reduce costs and increase throughput and scale by leveraging leading edge sequencing, sensor, and robotics technologies, leading to analysis by novel predictive computational tools. Together, these will accelerate plant breeding and increase the specificity of breeding targets.

The only way to design superior cultivars and improved breeding populations for crops grown under extreme environments is to extensively test them in the field. The ability to analyze and thus ascertain how genes ultimately determine traits in the field is limited by the national capacity to phenotype in the field. ARS scientists are key partners in the development of new automated field systems for measuring and analyzing phenotypes across many different crops.

Building on these advances, ARS researchers under this Problem Statement 1A and Problem Statements 1B and 2A will exploit crop genetic resources (older varieties, landraces, crop wild relatives, elite and experimental breeding stocks, and genetic mapping populations) to identify and understand the genetic bases for key traits. They will elucidate the genetic diversity and population structure of crops and crop wild relatives to reveal their potential for crop improvement. New research populations will also be generated, including mapping populations, mutant lines, and introgression populations. These materials will be characterized genetically by next-generation genotyping, phenotyping, and analytical approaches. The resulting genetic tools and insights will lead to testable hypotheses for novel plant breeding methods that maximize rates of genetic gain, and result in genetically enhanced populations improved for valuable agricultural traits, many of which are complex and quantitatively inherited. These traits will be analyzed and catalogued with unprecedented levels of detail, and the range of effects and frequencies of relevant genes will be functionally characterized in both improved and pre-breeding populations.

A thorough understanding of genome structure and genetic variation will enable more extensive trait analyses. Genetic and physical maps will be integrated, and comparative analyses of related genomes and shared genomic sequences will improve the accurate assembly of physical maps. Once advantageous alleles for specific traits are identified, genetic maps and reference genomes – highly saturated with genetic markers – will facilitate marker-assisted introgression to improve the traits. Introgression of existing variation from well characterized genetic resources and creation of new sequence variation will ensure a
sufficiently deep breeding pool to reduce genetic vulnerability and enable future genetic gain from selection. To exploit this new variation most effectively, new crossing and selection schemes to effectively recombine genomes and introgress useful alleles into new varieties will be developed.

**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. This knowledge will be provided by sequencing and annotating of plant genomes, and localizing and functionally characterizing genes affecting priority traits. 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 and efficient quantitative phenotyping technologies must be applied to analyze these populations.

High-resolution, sequence-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 of these alleles must be understood so that their effects can be more efficiently measured and influenced.

Innovative high throughput, quantitative phenotyping methods are needed because trait measurement is now often the critical bottleneck to genetic analysis and breeding progress. This must entail improving efficiency, precision, and cost-effectiveness of trait analyses through automation, increased throughput, or parallelization and use of new sensors (images, radiation wavelengths, chemistries, barcoding systems, radio frequency identification, and more).

Phenotypic, genetic, and other “-omic level” data from large populations must be combined in a coordinated effort to predict the performance of untested materials. All this must take into account the effects of the environment and management strategies on the expression and phenotypes of each trait. 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), different breeding system constraints, and a wide range of specialty and other crops, which until now have lacked these genetic tools.

**Anticipated Products and Accomplishments:**
- New traits developed via identification and analysis of existing genetic variation
• Well characterized genetic mapping populations, mutant lines, and other germplasm designed for trait discovery, analysis, and increasing genetic diversity in crop plants
• New variation in agriculturally important traits created through mutagenesis, biotechnology, genome editing, and/or other means
• Pipeline for accelerated trait discovery through functional genomic analysis and gene identification via genetic mapping
• New phenotyping approaches for important traits
• Methods to increase efficiency of genotyping
• Priority genes and genomes sequenced and functionally analyzed
• Improved methodology for predictive analyses that use specially-designed training populations, integrated “multi-omic data,” and biological knowledge
• Improved plant breeding techniques that apply knowledge from genetic interaction with environment and management factors
• New methods to effectively recombine genomes and introgress useful alleles and novel diversity

**Potential Benefits:**
The ultimate benefits from research conducted under Problem Statement 1A 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 combine disparate large data sets (also see Problem Statement 4A) 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 and full sequences, functionally characterized genes, and methods to combine data sets and recombine genomes) will elucidate the genetic basis and control of traits to 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. When feasible and applicable, genomic selection techniques and knowledge of gene interactions with each other or with the environment or crop management practices will be extended to specialty or other crops which have not benefitted from new tools to date. Improvements in and application of genome editing, mutagenic screens, and other genetic modification technologies will further enable creation of genetic variability and stocks to test hypotheses and accelerate cultivar improvement.
Problem Statement 1B: New crops, new varieties, and enhanced germplasm with superior traits.

To exploit rapid advances in technology, meet shifting consumer demands, and supply crop products to adequately support an increased population, U.S. agriculture requires new crops, new varieties, and enhanced germplasm with superior traits. The superior traits include higher yields, resistance or tolerance to biotic stresses and environmental extremes, improved agronomic or horticultural characteristics, and specific superior processing and product quality properties. Well-characterized genotypes enable plant breeders seamless access to materials needed for improving priority traits. Efficient breeding strategies are needed to utilize the available diversity, especially breeding with crop wild relatives. New technologies are required to enhance traditional methods of genetic improvement and breeding to genetically improve a broad spectrum of major, specialty, and new crops, including crops for alternative production systems. Genotyping and phenotyping germplasm and breeding lines for biotic resistance and abiotic tolerance traits will continue, but increased attention is needed for consumer-oriented quality traits to more adequately balance the needs of industry and the consumer in the breeding programs.

Although the past accomplishments of this National Program in breeding have been recognized, more effective breeding methods to mine the germplasm collections for economically important traits are needed. Also, adoption of high throughput genotyping and phenotyping technologies will be needed to enhance breeding efforts. Superior cultivars must be adapted to a wide variety of multiple cropping systems, and to competitive global market needs. The increased demand for enhanced germplasm will require breeders to utilize the improved genotyping, phenotyping, and analytical approaches that rely on accurate genetic and physical maps to facilitate novel selection methods. Faced with a shrinking supply of land, labor, water, and inputs, new cultivars and new crops developed by ARS will have more optimal plant architectures, enhanced water and input-use efficiencies, and optimal production efficiencies. New crops are also 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. In addition, a more concerted effort is needed during the next five years to breed U.S. crops adapted to environmental extremes and changing climates.

Research Needs:
Improved, well-characterized, and genetically diverse germplasm, breeding lines, and varieties are needed to produce consistently high yields of high quality crops under a variety of environmental conditions and agricultural production systems. Well maintained, available, and characterized genotypes in the National Plant Germplasm System are needed to enable breeders to quickly incorporate resistance to emerging diseases. More effective methods to mine the germplasm collections for economically important traits are needed. Adoption of high throughput genotyping and phenotyping technologies will enhance improvement efforts. Needs for improved resistance to biotic and abiotic stresses include more durable resistance and tolerance to multiple diseases and stresses. Varieties which are more efficient in their use of water and fertilizer are needed. Varieties with modified plant structures and architectures are needed to
maximize production efficiency and quality. Seeds and other propagules of new varieties must produce a uniform plant stand with adequate vigor under commercial production systems.

U.S. producers and consumers need improved crop product quality, including more flavorful and attractive fruits, vegetables, and ornamentals, more nutritious food and feed crops, prolonged product storage life and quality, and reduced energy requirements for processing. New crops and crop plants with new traits are needed. Crops are needed for new uses and new production systems, including protected environments such as greenhouses and tunnel systems, organic systems, small farms, home gardens, and cityscapes. Non-food crops capable of growth on underutilized agricultural lands must be bred as feedstocks for bioenergy production. There is a need to improve yield and product quality of crops with bio-based products, such as for bioenergy applications. Crop plants are needed that incorporate new value-added traits that improve end-use quality and broaden product utilization.

**Anticipated Products and Accomplishments:**
- Higher yielding genetically diverse crop plants
- Crop plants with resistance or tolerance to diseases and pests
- Crop plants tolerant to environmental changes or extremes
- Crop plants with enhanced water and input-use efficiencies
- Crop plants with architectures optimized for production efficiency
- Crop plants with improved seed/propagule quality for production systems
- Crop plants with superior product quality for consumers and producers
- Crop plants with enhanced nutritional quality
- Crop plants for economically viable bioenergy and bio-based products
- New crops and current crops with new traits for new uses

**Potential Benefits:**
Superior crop varieties and enhanced germplasm will contribute to a safe, plentiful, high quality food, feed, fiber, fuel, and ornamental supply with good price stability and a stronger, more secure agricultural industry. Crops with higher yields can help to stabilize rural economies, promote the 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, in addition to supporting better protection of the environment. Markets for reduced pesticide or pesticide-free agricultural products can be expanded. Crops with resistance or tolerance to multiple diseases, pests, and abiotic stresses 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 that utilize inputs more efficiently can reduce production costs and enable cultivation in marginal environments. Improved plant architectures contribute to increased high-quality yield per unit area, thereby enhancing the economic viability of small producers and family farms, as well as large-scale commercial production agriculture. 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.

Productive, efficient, non-food bioenergy crops adapted to marginal lands can 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 can provide alternative markets for farmers and a variety of novel products for domestic and international consumers. Fruits, vegetables, and herbs grown in non-traditional production settings, such as home gardens, patios, and rooftop gardens, can increase the amount of fresh produce available to urban consumers.

Component 2: Plant and microbial genetic resource and information management.

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

Food security and a strong agricultural economy will require access to genetic resources conserved in ARS genebank collections maintained by this National Program. Plant genebank collections operated by NP 301 as the National Plant Germplasm System (NPGS) provide germplasm, traits, associated information, research tools, and knowledge that contribute to achieving crops with higher yields, increased nutritional quality, abiotic and biotic stress resistance, and tolerance to environmental extremes, thereby contributing to resilient food and agricultural production systems. The NPGS maintains one of the most comprehensive plant genetic resource collections in the world with 576,000 accessions representing more than 15,000 plant species that are important to agriculture. NP 301 curators and researchers also conserve, characterize, and distribute more than 100,000 different cultures of pathogenic or beneficial microbes that are important to crop production, industrial processes, food safety, and human health. It is a priority for NP 301 to ensure that its ex situ genebank collections have high levels of genetic integrity, health, and viability and are maintained in secure modern facilities. The positive impacts of the NPGS and microbial genetic resource collections are demonstrated in part through the distribution of an average of 250,000 plant accessions per year – the highest distribution rate of any plant genebank system in the world.

Genebank curators are challenged with not only meeting the needs of current customers, but also anticipating the genetic resource needs of the future. This requires creativity and innovation to maintain and strategically expand plant collections to include more crop wild relatives, landraces, and cultivars of historical importance. Similarly, maintenance and strategic expansion of collections are needed for microbial type strains, microbial germplasm required to
advance national research priorities, and to safeguard “orphaned collections” of significance for industrial processes, crop productivity, and combating crop diseases. Close interactions with customers will ensure that the characterization and evaluation of genetic resources generate data that are critical for effective acquisition and curation and will facilitate the identification of optimal germplasm for research, education, and breeding. National and international collaborations are essential for sharing genetic resources and information, conducting research, combating genetic erosion, and strengthening collections according to the standards of access and benefit sharing established by both domestic legislation and international agreements.

Future agricultural advances are dependent on having open and free access to data. Information associated with NPGS collections is stored, curated and made publicly available via GRIN-Global, an advanced information management system that is readily accessed by more than 1.5 million unique “visits” annually (as indicated by IP addresses). GRIN-Global will become even more valuable when it is interoperable with other public databases that maintain additional environmental, geospatial, pedigree, phenotypic, and genomic information. Data about microbial germplasm are maintained and delivered through databases operated by individual NP 301 microbial genebanks.

**Research Needs:**

Agricultural systems throughout the world have many needs for NP 301’s plant and microbial genetic resources and associated information. High quality plant and microbial genetic resources must be safeguarded in state of the art genebanks and well documented in GRIN-Global or microbial germplasm databases. Collection best management practices must be updated, developed or improved, and applied as consistently as possible to ensure the efficient, effective, long-term, and safe preservation of critically important genetic resources. In-house (ARS) and collaborative research supporting genetic resources preservation and effective use should address priority areas, and results must be implemented and broadly disseminated. GRIN-Global and microbial germplasm databases must continue to evolve and will require improvements and customizations to ensure an easy-to-use interface supporting genebank curation. GRIN-Global and microbial germplasm databases must also continue to provide customer and stakeholder access to passport, phenotypic, and genotypic data to make the best use of genebank resources.

Growth of genebank collections usually outpaces budgets and available space, requiring genebanks to perform research to find more efficient and effective methods to conserve genetic resources. Collections must be protected from pathogens, pests, and environmental extremes, and this will require development and adoption of novel and creative science-based practices. Priorities for new acquisitions and management of existing genebank collections must ensure genetic integrity of the conservation target (i.e., genes or genotypes), allowing flexible strategies depending on the life history and reproductive biology of the species and improvement status of the accession (i.e., elite line vs. wild relative, presence of genetically engineered traits). Collection viability and genetic integrity must be maintained for the long-term by improving collection
management practices and implementing regeneration and storage strategies that minimize genetic drift and lengthen the intervals between regeneration events. Secure back-ups are needed for all materials, and plans must be made to attain this goal, particularly for vegetatively propagated and non-orthodox seeded specialty crops for which cryobiological technologies must be developed and implemented strategically and judiciously.

Exchanges and plant explorations must acquire well-documented germplasm. Research is needed to identify the genetic coverage of collections through approaches that consider taxonomy, geospatial distribution, genetic diversity and structure, and phylogenetic relationships. When gaps are identified and prioritized, they must be filled through appropriate methods to capture the desired conservation target, especially for priority microbial germplasm such as type strains, “orphaned collections” significant to agriculture and industrial processes, crop wild relatives, landraces, and important historic cultivars. Crop wild relative (CWR) inventories should be compiled as one of many tools available for prioritizing collection efforts for wild species. Domestic collaborations and partnerships must be developed to secure priority CWR species in situ. Important historical cultivars and genetic stocks can be identified from archived documents and through consultations with crop research communities and Crop Germplasm Committees (CGCs). International collecting efforts must support appropriate benefit sharing. GRIN-Global and microbial germplasm databases must incorporate data that document the collecting environment and enable geospatial tools that can help identify germplasm better adapted to abiotic or biotic threats.

Data generated or collected must take into account the critical needs of stakeholders by recording them with prioritized descriptor and standardized trait/ontology lists for phenotypic evaluation and uniform methods for genotypic characterization of plant genetic resources. Different descriptors might be needed depending on the improvement status of accessions (wild vs. cultivar). Partnerships must be established to generate and deliver germplasm evaluation data to meet new challenges, such as climate change. Evaluations must be conducted under multi-environment, multi-production systems. Future evaluations must address difficult to evaluate traits (e.g., roots) and new emphases (e.g., quality traits such as nutritional content).

High throughput phenotyping of plant genebank accessions must be coordinated with genotyping by sequencing and genetic marker analyses. Active collaborations must be pursued with other NP 301 Action Plan Components; other ARS National Programs, particularly NP 107, NP 303, NP 304, NP 305, and NP 306; multi-institutional genomic research consortia; other genebank systems; and with new partners to facilitate cross-dataset comparisons. Outcomes from phenotypic, genotypic, genomic, and diversity assessments must be incorporated into curatorial decisions regarding increasing collections’ genetic diversity, monitoring integrity, and reducing redundancy. Ideally, genetic diversity assessments would be conducted either before or soon after acquisition to identify efficient management strategies for wild-collected and elite
germplasm. Core and special purpose subsets should be defined. Comprehensive germplasm evaluations must be conducted, and the resultant data incorporated into GRIN-Global.

The NPGS must supply genetic resources to the domestic and international scientific and educational communities according to national policies and the requirements of the FAO International Treaty on Plant Genetic Resources for Food and Agriculture. Plant and microbial genetic resources must be high quality, healthy, viable, true-to-type, well-documented, and distributed to customers in a timely manner. Customers must continue to receive excellent service regarding both distribution and scientific interactions. GRIN-Global and microbial germplasm databases must be kept up-to-date with taxonomic, passport, environmental, geospatial, phenotypic (trait evaluation), genetic, inventory, and availability data so customers can refine their orders and focus their research.

Genetic resource management capacities must be strengthened by cooperation with national and international public and private sector institutions and enterprises to support mutually beneficial activities that include germplasm characterization, evaluation, regeneration, exchange, and crop improvement. Partners include stakeholders, CGCs, Regional Technical Advisory Committees, students, scientists, agricultural producers, and other customers involved in outreach activities. CGCs offer a wealth of knowledge from experts in crop communities, and curators must actively seek the involvement of CGCs when developing and updating Crop Vulnerability Statements and other management tools to ensure collections continue to be valued and effectively used by customers. National and international research, breeding, and educational partnerships must be strengthened to promote technology transfer, innovation, and conservation of germplasm for the long-term. GRIN-Global and microbial germplasm databases must continue to be expanded and updated to streamline data entry and to comprehensively document passport, phenotypic, and genetic data to meet the curatorial needs of all accession types (e.g., clonally-propagated samples). Information management capacities must be strengthened by developing practices for standardizing, archiving, and retrieving data in ways that will enhance and extend the functionality of GRIN-Global through database interoperability.

**Anticipated Products and Accomplishments:**

- High quality plant and microbial genetic resources and associated information safeguarded in state-of-the-art genebanks and information management systems
- More efficient and effective genetic resource and information management approaches and practices developed
- Gaps in genetic resource collections identified and priority genetic resources (especially crop wild relatives) acquired to fill those gaps
- Genetic resource characterizations and key trait evaluations conducted and data incorporated into information management systems
• High quality genetic resources and associated information delivered to customers
• Genetic resource and information management capacities strengthened through partnerships

**Potential Benefits:**
The products anticipated from the preceding genetic resource management operations and research will improve the quality of national and international genebank collections and ensure continued worldwide access to critical genetic resources. Secure collections will be high-quality, well-documented, and have excellent vigor, viability, and health. Improved management efficiencies and effectiveness will be achieved as new genetic resource management approaches and collection best management practices are updated, developed, and implemented. Strengthened partnerships with national and international organizations will enhance the ability of NPGS to effectively carry out its mission. Crop Vulnerability Statements and other collection assessments will ensure that the NPGS continues to safeguard critical genetic resources and meet the needs of customers. Collections will be strategically expanded through explorations and exchanges. Standardized genotypic and phenotypic diversity data will facilitate multi-crop comparisons. Expanded evaluations will provide information for priority agricultural traits to researchers, breeders, and producers. More requests for genetic resources will be better targeted and specific, thus increasing the impact of distributions without unnecessarily depleting the available materials.

Accurate and comprehensive genetic resource information will be made available to the research and education communities. Functionality of GRIN-Global will be expanded as it becomes interoperable with external public databases that contain relevant associated data. GRIN-Global enhancements will provide customers with the tools they need to target specific accessions for requests. The efficiency of genetic resource utilization will increase as relevant, timely, high-quality comprehensive characterization and agricultural trait data are available to domestic and international communities. The world-class genetic resources collections maintained by ARS will safeguard and provide access to genetic resources, key crop or microbial traits, associated information, tools, and knowledge that will ultimately benefit farmers and consumers on a global scale.

**Component 3: Crop Biological and Molecular Processes**

ARS scientists conduct fundamental – often long-term and high-risk – research to generate an enhanced knowledge base and to 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. Research conducted under this component will elucidate the biological processes underlying crop
productivity and quality, and develop new means for assessing the potential effects and 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 will 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 fundamental investigations with more applied methods to develop focused strategies for solving specific problems. Specific areas for research 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’ fundamental plant research. It includes not only discovery research, but also efforts to improve biotechnology risk assessment methods and to analyze potential unintended consequences of genetic engineering and transgenes for crop plants, for crop production, and for the environment. This involves developing superior new public domain genetic tools that improve the efficiency of conventional plant breeding and that also assess the potential risks associated with genetically engineered crops.

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

Crop plants provide food, feed, fiber, fuel, and ornamentals as well as primary and secondary metabolites that are beneficial to health, nutrition, flavor, and industrial applications. Agricultural production must increase by sixty percent in the next fifteen years to meet global demand. Although crop production must increase, fewer land, water, and nutrient inputs will be available. Disease, pests and climate variability will further constrain agricultural production. Traditional plant breeding approaches alone will be unable to meet these challenges. To aid in plant biotechnology and marker assisted selection, our fundamental understanding of the molecular, physiological, and genetic bases of agronomic and horticultural performance must improve.

To be successful crops, plants must efficiently convert light, water, and nutrients into high-quality and high-value plant organs and tissues. Crops must function in terms of optimal patterns of architecture, development, and biochemistry, in a given environment. At the same
time, crop plants continually sense and actively respond to thousands of different environmental cues. These include changes in light or air quality, high or low temperature stress, nutrient and water availability and presence of beneficial and pathogenic microbes above and below ground. The effectiveness of the plant response varies greatly throughout development, among varieties, and across geographical regions. 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 and development, and environmental responses.

To meet the needs of consumers and producers, ARS research will exploit the genetic variability in plant and microbial germplasm collections to identify new traits and the underlying genes and genetic networks that govern those traits to produce crops with improved plant performance across diverse environments. This will require highly coordinated and integrated research in the field, the lab and across cyber infrastructure. ARS research will apply expertise in plant phenotyping, genomics, proteomics, metabolomics, epigenetics, metagenomics, and bioinformatics to identify the basic mechanisms regulating agriculturally important traits. Improving our understanding of these processes will lead to crop varieties with improved productivity, quality, value, and sustainability.

**Research Needs:**
To better understand molecular, genetic, and physiological processes underpinning important crop traits, the capacities to characterize and measure complex plant traits and collect, share and store data across crop species must be expanded. Gene-editing technologies, next-generation sequencing, and proteomic and metabolomic techniques must be applied according to rigorous scientific principles in order to generate datasets for global gene and protein expression or metabolite levels that enable systems level comparisons both within and across crop species. Molecular, biochemical, and physiological phenotypes must be linked 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. These approaches must be combined with the development of new tools to test gene function and development of new algorithms and computational tools for data analysis, integration, and sharing.

A more complete and detailed understanding of plant growth and developmental regulatory processes affecting food quantity and quality is necessary for developing plant breeding and biotechnology-based solutions for maintaining and increasing crop productivity in response to a changing climate. The genes and networks underlying crop adaptation to biotic and abiotic stress and multi-species interactions within the microbiome must be identified. Genetic networks must be explored and exploited across the diversity present within our germplasm collections. The potential for both identification of superior alleles and gene modification must be assessed for its role in altering the fundamental biological processes underlying improved crop productivity, superior quality, and enhanced resource use efficiency.
In addition, developing a broader understanding of the plant microbiome and its contribution to crop health and disease will be key for meeting agricultural yield goals in the coming decades. This understanding must include characterization of the role of these microorganisms in modulating host fitness at the molecular and biochemical level, and a refined knowledge of the environmental and host-mediated factors that influence community recruitment and development.

**Anticipated Products and Accomplishments:**
- Enhanced knowledge of existing diversity in crop plant interactions with abiotic and biotic environmental factors at the molecular, whole genome, and systems levels
- Enhanced knowledge of how plant microbiomes positively and negatively influence crop performance at the genetic, molecular and physiological levels
- Enhanced knowledge of how growth and development of crop plants are controlled at the genetic, molecular, and physiological levels
- Enhanced understanding of the biochemical pathways and metabolic processes that underpin crop traits and improvement
- Enhanced knowledge of food (seed, fruit, tubers, etc.) qualities and nutritional value at genetic, molecular and physiological levels
- New biotechnological tools to test and transfer gene function for improved crop varieties
- New tools for applying fundamental knowledge to crop improvement

**Potential Benefits:**
Knowledge of the molecular mechanisms underlying crop adaptation and plant-environment interactions will lead to greater crop productivity in unfavorable growth environments and improve the resilience 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, optimal plant architecture, and reproductive success. Understanding the molecular mechanisms controlling key developmental transitions in plants 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 NP 301 scientists to improve crop plants for adaptation to changing climates, increased carbon sequestration, and value-added production. Knowledge of plant metabolism combined with understanding of plant growth and developmental processes will enable optimization of biochemical pathways for synthesizing valuable chemical products. Knowledge of the plant microbiome will improve crop productivity and tolerance of environmental stress through identification of beneficial microorganisms with the ability to increase nutrient availability, improve defense response, and modulate crop growth and development.
Problem Statement 3B: Crop biotechnology risk assessment and coexistence strategies.

Genetic engineering offers tremendous promise for improving crop yields and protection by increasing the sustainability and reducing the environmental footprint of production systems. The products derived from these technologies include foods with improved quality and nutritional properties, as well as crop plants with enhanced disease and pest resistance and improved agronomic traits. Research that integrates the development of efficient biotechnological improvement technologies combined with the study of their potential effects on agricultural production and the environment will help advance the application of these technologies. In addition, except for a relatively few commodities, methods for efficient crop genetic engineering have not been fully developed. Additional techniques for broadly and efficiently applying these technologies to a wide range of crops and their constituent varieties are required.

Crop genetic engineering includes the production of plants with gene sequences from diverse sources, arranged in unique combinations with novel patterns of expression. Although this technology offers great promise for crop improvement, a careful evaluation of the potential effects and unintended consequences of these products is necessary to facilitate the coexistence of conventional and biotech crops in global agriculture. In addition, this research will provide critical support to regulatory agencies charged with ensuring the safe utilization of biotech crops and the new technologies that are being developed. The intersection of genetic engineering, genomics, and plant breeding approaches have, and will continue to yield novel technologies for the modification of plant genomes and the generation of improved crops. Research is required to provide a technical foundation for addressing the potential concerns of unintended impacts of crop genetic engineering and removing barriers to coexistence of different crop production systems.

Research Needs:

New techniques and approaches for the evaluation and mitigation of unintended consequences of biotechnological improvements on crop plants, crop performance, and the environment are needed. Such approaches include methods and molecular tools for the precise genetic modification of crops to enable increased specificity and limit off-target effects, as well as for post-transformation removal of unnecessary DNA sequences and the characterization of off-target modifications. Techniques are needed for enabling the “pyramiding” of multiple transgenes and/or gene edits for the introduction of multigenic traits. Also, biotechnology methods with broader applications are needed to benefit a broader array of agriculturally significant crop species, both of major commodities and specialty crops.

It is important to emphasize that although genetic engineering technologies have revolutionized the production of a few crops, the absence of specific enabling technologies, as well as other factors, has prevented its application to the vast majority of crops. New and innovative approaches must be developed for the biotechnological
improvement of diverse crop species and varieties, including improved methods of transformation and regeneration, to enable routine, rapid and efficient deployment of varieties improved by both standard genetic engineering and novel genome modification technologies. As more crops become genetically engineered, associated methods to evaluate their potential risks and to meet regulatory requirements must be refined.

Anticipated Products and Accomplishments:
- Improved public domain tools for producing genetically engineered crops
- Improved methods for application of biotechnology to a broader selection of crop species/varieties
- Genetic engineering technologies that improve the efficiency of conventional crop breeding
- Methods to identify and reduce unintended effects of biotechnological improvement on crop plants, agricultural production, and the environment

Potential Benefits:
Improved public domain tools and strategies for producing and characterizing genetically engineered crops will enable more rapid and safe deployment of improved commodities and specialty crops. The identification and reduction of unintended effects of these biotechnological improvements will ensure protection of the environment and facilitate the coexistence of agricultural markets. Collectively, these efforts will promote the integration of genetic engineering, molecular breeding, and crop biological knowledge, improving the capacity to generate and deliver crops that can address the future challenges facing U.S. agricultural production.

Component 4: Information resources and tools for crop genetics, genomics, and genetic improvement.

Problem Statement 4A: Information resources and tools for crop genetics, genomics, and genetic improvement.

Breeding, genetics, genomics, and genetic resource management (see Components 1, 2, and 3 of this Action Plan) are information sciences. Innovative genomic approaches and analytical capacities have produced an amazing information catalog of the genetic and molecular diversity of our crops. The capacity to collect data for field performance is being significantly amplified by numerous automated measurement systems. The transformation of these data to knowledge which is applicable to breeding, genetics, and genetic resource management will require a proportionately strong bioinformatic infrastructure. Improved bioinformatic tools and services built on a powerful computational infrastructure are needed to handle changing data types and increased volumes of data and scales of datasets — sometimes a thousand to a million fold — that will present tremendous opportunities for more rapid crop genetic improvement.
During the last five years, genomics-assisted breeding, which integrates data generated by genomic analyses with yield data for more rapid selection of superior breeding lines, became a standard procedure in private-sector breeding programs for some major crops. However, for many crops, the data and knowledge infrastructure are insufficient for exploiting these new analytical tools. To advance to the next level of crop improvement where natural genetic variation and that generated by precise genome editing are more generally available, “biology-assisted breeding” is needed, especially to improve crop product quality and yield. Designing our next generation of crops will be based on genomic information integrated with information from genetics, physiology, molecular biology, biochemistry, environmental interactions, and evolutionary diversity. NP 301 research will continue to lead the integration of these disparate, commonly non-standardized data streams to facilitate more efficient crop genetic improvement and the long term sustainability of our crop production systems.

Although more effective information integration methods are key, data collection costs are decreasing so rapidly that “biology-assisted breeding” will be accessible, not only to major crops, but also to specialty crops and forages. Working across ARS National Programs and with international partners, NP 301 will establish superior software modules to manage and analyze numerous biological data types relevant to crop improvement. It will also lead in developing data exchange protocols that facilitate access, analysis, and integration of these data sources. Building on these principles, NP 301 research will deliver bioinformatic solutions for the full range of crops important to U.S. agriculture.

**Research Needs:**
Crop improvement has traditionally depended on farmer- or breeder-selections of superior varieties from relatively small populations. This process has been remarkably successful for realizing dramatic initial genetic gains and for selecting large-effect phenotypic traits. But it is increasingly difficult to make substantial genetic gains within the current elite crop varietal gene pool. Improving elite varieties, or making rapid improvements for any crop, requires much more intensive exploitation of information about many factors: available crop diversity, performance of genotypes in multiple environments, accurate characterization of plant structure and agronomic and horticultural traits (phenotypes), genetic markers for priority traits, gene and regulatory functions, and lengthy genome sequences (especially haplotypes) for ordering, anchoring and integrating genomic features valuable for crop breeding.

The information sources needed to support crop genetic improvement can be categorized according to several, mostly distinct dimensions:
- a germplasm dimension, with information about stocks held in a breeding program or in the National Plant Germplasm System;
- a genotype dimension, with information about genetic variation across genepools;
- a genome sequence dimension, including positional information for genes and other genetic features; and
• a phenotype dimension, with information about agricultural traits for any breeding line or germplasm accession.

Managing the preceding multi-dimensional information for crop improvement requires databases to hold the complex data types mentioned above, as well as intuitive interfaces to those databases to enable researchers to efficiently access that information. Data standards and metadata formats are needed to enable data curators to efficiently collect, organize, store, and access diverse and complex biological data. Visualization and analysis tools are essential for enabling this information to accelerate the progress of plant breeding and research.

Information across these dimensions must be integrated to identify optimal elite breeding lines, and to locate germplasm accessions that are either genetically identical or genetically distinct from elite lines in targeted genomic regions. This kind of analysis involves querying and visualization tools from several databases and portals, and interlinkages, such as, information about germplasm (via GRIN-Global); information about traits and quantitative trait loci, proteins and metabolic pathways, genes and genomic frameworks from species or clade databases (e.g., SoyBase, Gramene, GrainGenes, MaizeGDB, or the Triticeae Toolbox); and genotypic data from species databases or from a multi-species genotype database and tool collection sets, such as TASSEL or GOBII.

To accommodate data volumes and complexities and to support the data needs of more crops, the more extensive computational resources of SCINet (ARS), CyVerse (NSF), XSEDE (NSF), or public “cloud” data storage capacity, such as Amazon Web Services (AWS), are needed. Overall ARS data management solutions must ensure the continued maintenance and support of databases along with sufficient IT infrastructure for data transfer, data storage, and computational capabilities.

ARS database projects will increasingly require adoption of best practices for software development, documentation, and sharing — including reuse and extension of open-source software programs — common database schemas where feasible, and web APIs (application programming interfaces) to facilitate data access by other projects. For efficient data collection, analysis, and sharing, NP 301 scientists must adopt and promote metadata standards suitable for heterogeneous data types; participate in refining common data standards, formats, and protocols for data collection and storage; coordinate with other plant database projects inside and outside ARS; and apply curatorial and development expertise from many groups. Those approaches are needed to maintain the flexibility to create and test new interconnected, modular, data-driven software tools according to a rapid development cycle.

Across ARS research programs, the rapid growth of information presents opportunities for significant advances in crop production. Nonetheless, those same opportunities represent substantial challenges for data management, including physical data storage,
organized file storage, and searchable and retrievable data for future analysis by pertinent databases. An effective ARS data management system is needed for standardizing data ingestion, data preservation, and archiving, to make ARS research operations more efficient. In addition, Federal policy requires that scientific data generated by Federal funding are retained and accessible.

**Anticipated Products and Accomplishments:**
- Interconnected and searchable information resources and tools actively and efficiently curated and maintained for crop research and breeding
- Information tools that connect genetic, genomic, and metabolic traits with specific germplasm accessions or samples
- New bioinformatic tools developed for data analysis and mining, and to manage high throughput phenotypic and genotypic data and knowledge
- Strategic partnerships with user communities and stakeholders maintained and strengthened

**Potential Benefits:**
The availability of integrated datasets and digital objects that can be navigated and analyzed with ease will enable scientists to generate and associate new knowledge about plant 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 facilitating data access and broadening participation in crop research. Such improvements will promote efficiencies in both time and cost that accelerate crop genetic research, breeding, and genetic resource management.

**NP 301 Resources:**

**Arizona**
Maricopa, AZ - U.S. Arid Land Agricultural Research Center, Plant Physiology and Genetics Research Unit

**Arkansas**
Stuttgart, AR - Dale Bumpers National Rice Research Center

**California**
Albany, CA - Plant Gene Expression Center
Albany, CA - Western Regional Research Center, Crop Improvement and Genetics Research Unit
Davis, CA - Crops Pathology and Genetics Research Unit
Davis, CA - National Clonal Germplasm Repository - Tree Fruit & Nut Crops & Grapes
Parlier, CA - San Joaquin Valley Agricultural Sciences Center - Crop Diseases, Pests and Genetics Research Unit
Riverside, CA - National Clonal Germplasm Repository for Citrus
Riverside, CA - U.S. Salinity Laboratory, Water Reuse and Remediation Research Unit
Salinas, CA - Crop Improvement and Protection Research Unit
Colorado
Fort Collins, CO - Center for Agricultural Resources Research, National Laboratory for Genetic Resources Preservation, Plant and Animal Genetic Resources Preservation Unit
Fort Collins, CO - Center for Agricultural Resources Research, National Laboratory for Genetic Resources Preservation, Plant Germplasm Preservation Research Unit
Fort Collins, CO - Center for Agricultural Resources Research, Soil Management and Sugarbeet Research Unit

District of Columbia
Washington, DC - U.S. National Arboretum, Floral and Nursery Plants Research Unit
Washington, DC - U.S. National Arboretum, Gardens Unit

Florida
Canal Point, FL - Sugarcane Production Research Station
Fort Pierce, FL - U.S. Horticultural Research Laboratory, Subtropical Insects and Horticulture Research Unit
Gainesville, FL - Center for Medical, Agricultural and Veterinary Entomology Chemistry Research Unit
Miami, FL - Subtropical Horticulture Research Unit

Georgia
Byron, GA - Fruit and Tree-Nut Research Unit
Dawson, GA - National Peanut Research Laboratory
Griffin, GA - Plant Genetic Resources Conservation Research Unit
Tifton, GA - Crop Genetics and Breeding Research Unit
Tifton, GA - Crop Protection and Management Research Unit

Hawaii
Hilo, HI - Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center, Tropical Plant Genetic Resources and Disease Research Unit

Idaho
Aberdeen, ID - Small Grains and Potato Germplasm Research Unit
Kimberly, ID - Northwest Irrigation and Soils Research Laboratory

Illinois
Peoria, IL - National Center for Agricultural Utilization Research, Bio-Oils Research Unit
Peoria, IL - National Center for Agricultural Utilization Research, Crop Bioprotection Research Unit
Peoria, IL - National Center for Agricultural Utilization Research, Mycotoxin Prevention and Applied Microbiology Research Unit
Urbana, IL - Global Change and Photosynthesis Research Unit
Urbana, IL - Soybean/Maize Germplasm, Pathology, and Genetics Research Unit
Indiana
West Lafayette, IN - Crop Production and Pest Control Research Unit

Iowa
Ames, IA - Corn Insects and Crop Genetics Research Unit
Ames, IA - Plant Introduction Research Unit

Kansas
Manhattan, KS - Center for Grain and Animal Health Research, Hard Winter Wheat Genetics Research Unit

Louisiana
Houma, LA - Sugarcane Research Unit
New Orleans, LA - Southern Regional Research Center, Cotton Fiber Bioscience Research Unit

Maryland
Beltsville, MD - Beltsville Agricultural Research Center, Genetic Improvement for Fruits & Vegetables Laboratory
Beltsville, MD - Beltsville Agricultural Research Center, Molecular Plant Pathology Laboratory
Beltsville, MD - Beltsville Agricultural Research Center, National Germplasm Resources Laboratory
Beltsville, MD - Beltsville Agricultural Research Center, Soybean Genomics & Improvement Laboratory
Beltsville, MD - Beltsville Agricultural Research Center, Sustainable Agricultural Systems Laboratory
Beltsville, MD - Beltsville Agricultural Research Center, Sustainable Perennial Crops Laboratory
Beltsville, MD - Beltsville Agricultural Research Center, Systematic Mycology & Microbiology Laboratory

Michigan
East Lansing, MI - Sugarbeet and Bean Research Unit

Minnesota
St. Paul, MN - Plant Science Research Unit

Mississippi
Mississippi State, MS - Crop Science Research Laboratory, Corn Host Plant Resistance Research Unit
Mississippi State, MS - Crop Science Research Laboratory, Genetics and Sustainable Agriculture Research Unit
Poplarville, MS - Southern Horticultural Research Laboratory
Stoneville, MS - Crop Genetics Research Unit
Stoneville, MS - Genomics and Bioinformatics Research Unit
Missouri
Columbia, MO - Plant Genetics Research Unit

Nebraska
Lincoln, NE - Wheat, Sorghum and Forage Research Unit

New York
Geneva, NY - Grape Genetics Research Unit
Geneva, NY - Plant Genetic Resources Research Unit
Ithaca, NY - Robert W. Holley Center for Agriculture & Health - Plant, Soil and Nutrition Research Unit

North Carolina
Raleigh, NC - Plant Science Research Unit
Raleigh, NC - Soybean and Nitrogen Fixation Research Unit

North Dakota
Fargo, ND - Red River Valley Agricultural Research Center, Cereal Crops Research Unit
Fargo, ND - Red River Valley Agricultural Research Center, Sugarbeet and Potato Research Unit
Fargo, ND - Red River Valley Agricultural Research Center, Sunflower and Plant Biology Research Unit

Oklahoma
Stillwater, OK - Wheat, Peanut, and Other Field Crops Research Unit

Oregon
Corvallis, OR - Forage Seed and Cereal Research Unit
Corvallis, OR - Horticultural Crops Research Unit
Corvallis, OR - National Clonal Germplasm Repository

Puerto Rico
Mayagüez, PR - Tropical Agriculture Research Station, Tropical Crops and Germplasm Research Unit

South Carolina
Charleston, SC - U.S. Vegetable Laboratory
Florence, SC - Coastal Plain Soil, Water and Plant Conservation Research Center

Texas
College Station, TX - Southern Plains Agricultural Research Center, Crop Germplasm Research Unit
Lubbock, TX - Cropping Systems Research Laboratory, Plant Stress and Germplasm Development Research Unit
**Washington**
Pullman, WA - Grain Legume Genetics and Physiology Research Unit
Pullman, WA - Plant Germplasm Introduction and Testing Research Unit
Pullman, WA - Wheat Health, Genetics, and Quality Research Unit
Wapato, WA - Temperate Tree Fruit and Vegetable Research Unit

**West Virginia**
Kearneysville, WV - Appalachian Fruit Research Laboratory - Innovative Fruit Production, Improvement and Protection Unit

**Wisconsin**
Madison, WI - Cereal Crops Research Unit
Madison, WI - Vegetable Crops Research Unit