

Research Challenges and Resource Needs in Cyberinfrastructure & Bioinformatics: **BIG DATA in Plant Genomics**

The National Plant Genome Initiative



Interagency Working Group on Plant Genomics

Diane Jofuku Okamuro

National Science Foundation, Biological Sciences Directorate

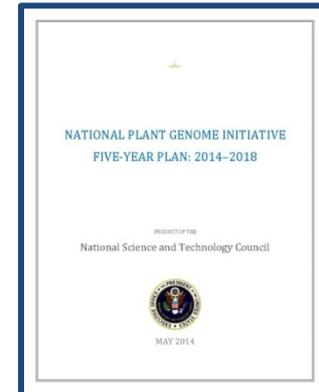
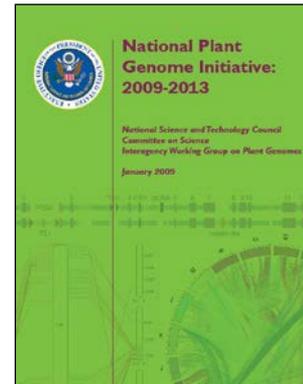
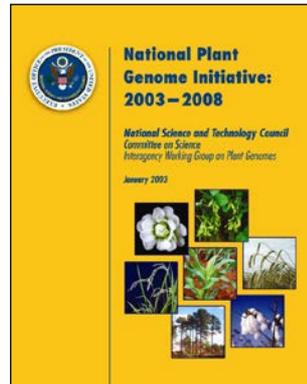
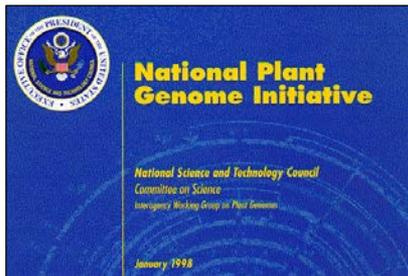
National Plant Genome Initiative (NPGI)

- Established in 1998 as a coordinated Federal program in the genomics of economically important crop plants
 - As part of the NPGI, established the **Plant Genome Research Program** (PGRP) – major focus on plants of agricultural importance and plant processes of potential agronomic value
 - Project activities expected to be on a genome-wide scale
 - Supports projects at all scales from single-investigator projects through multi-investigator, multi-institution projects
 - Research areas include structural genomics, functional genomics, bioinformatics, databases and tool/resource development

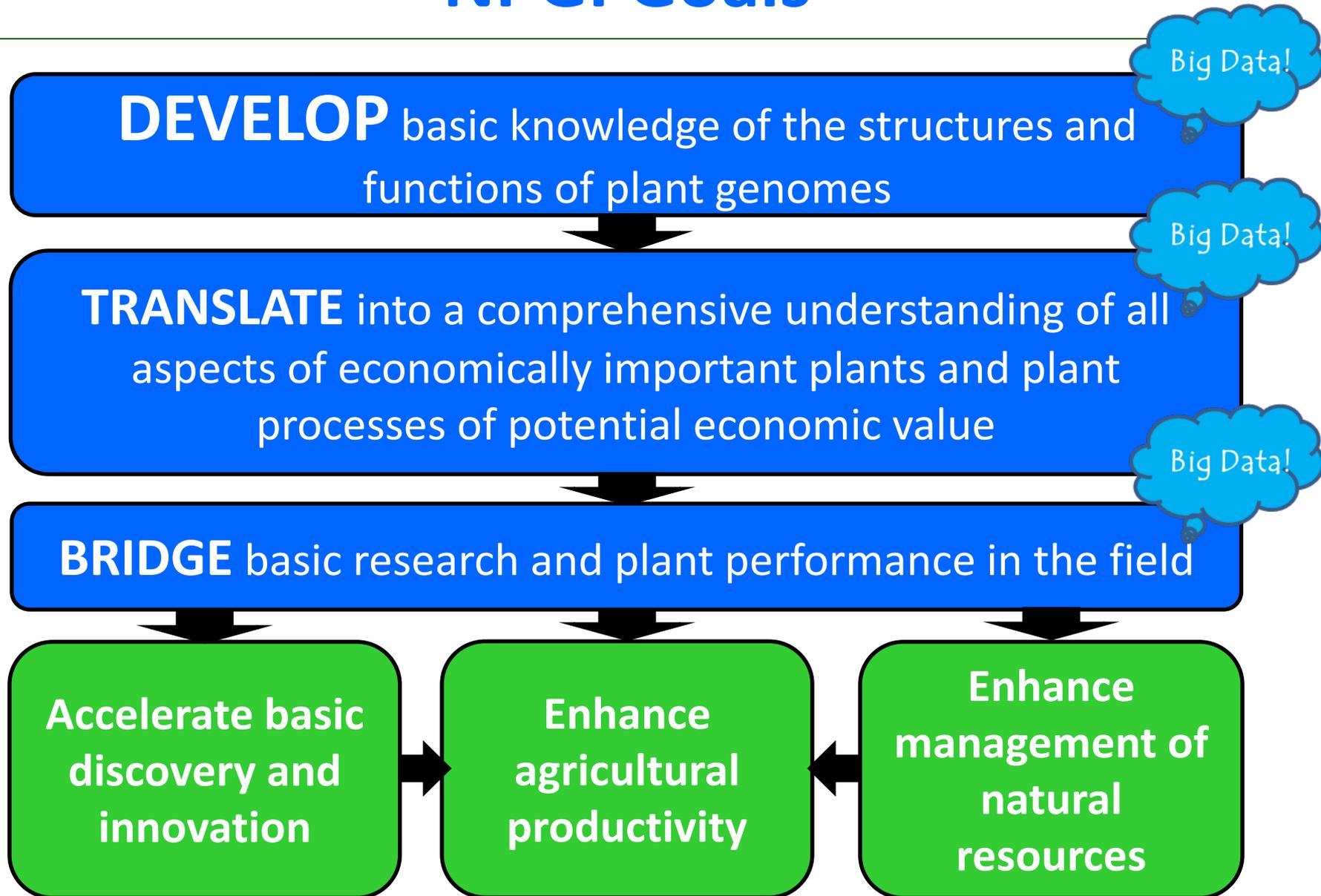


National Plant Genome Initiative (NPGI)

- Managed by the Interagency Working Group on Plant Genomics (IWGPG)
 - Current participants include NSF, US Department of Agriculture (USDA), US Department of Energy (DOE), National Institutes of Health (NIH), US Agency for International Development (USAID), U.S. Geological Survey, Smithsonian Institution, and Environmental Protection Agency (EPA)
- Activities coordinated through Five-Year Plans that are developed with community input and guided by principles first articulated in 1998 that include the **timely access to all outcomes of NPGI investments**



NPGI Goals

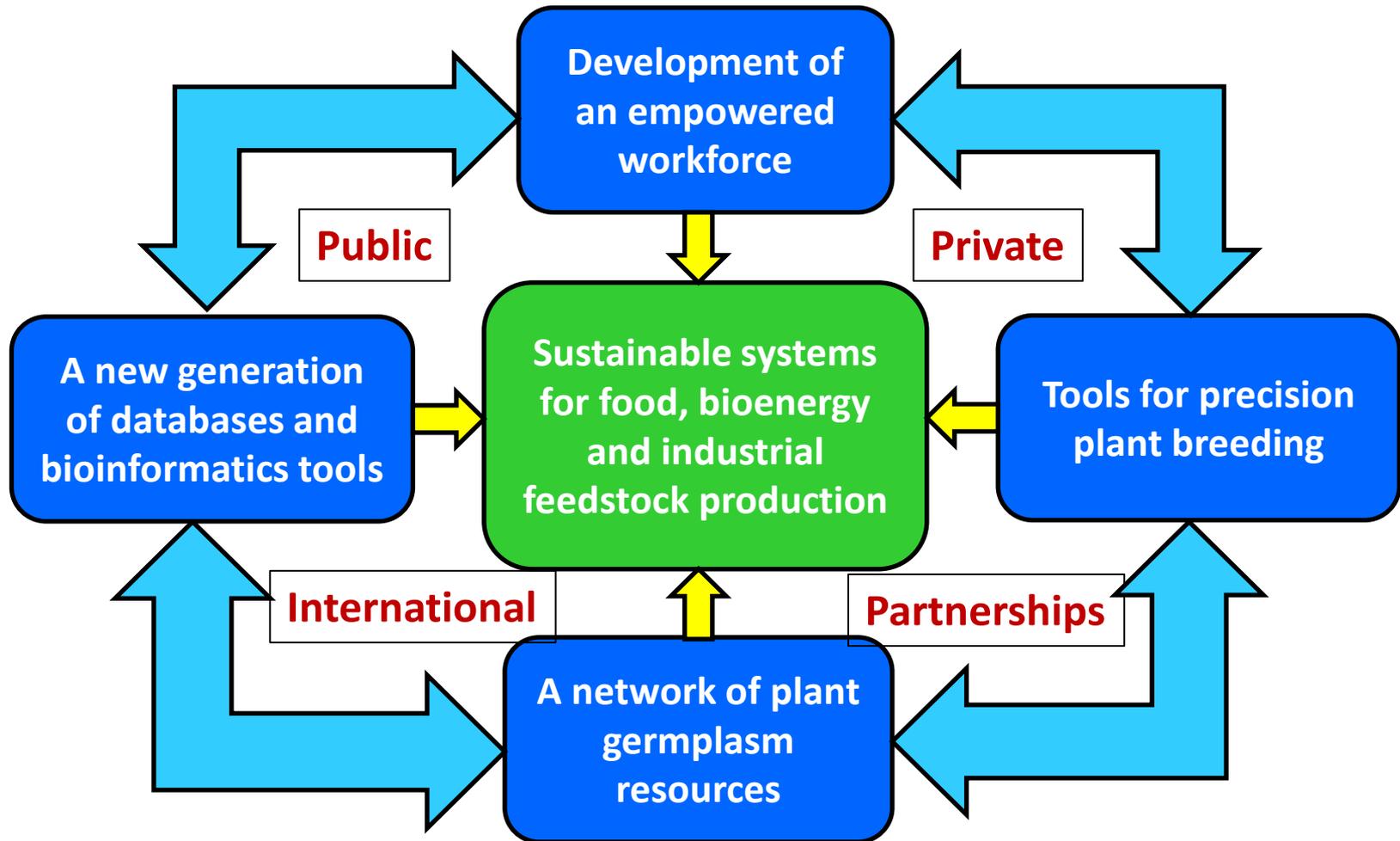


Objectives for 2014-2018

IWGPG chartered in 2014 to implement the goals set out in the *NPGI Five-Year Plan: 2014-2018*

- **Objective 1:** Develop new generation of databases and tools to enable basic discovery through translation to the field
- **Objective 2:** Create a network of plant germplasm resources to meet the Nation's research and breeding needs for the 21st Century
- **Objective 3:** Build tools to advance our understanding of plant biology and translate this knowledge to precision plant breeding for the development of sustainable systems for food, bioenergy, and industrial feedstock production
- **Objective 4:** Empower the workforce to use a new generation of tools and resources
- **Objective 5:** Build public-private partnerships to advance the translation of basic discoveries to the field and support innovation in the practice of agriculture
- **Objective 6:** Strengthen international partnerships to bring the benefits of new discoveries to all

NPGI Plan for 2014-2018





IWGPG Charter

IWGPG Chartered in 2014 to engage the plant science community in implementing the goals set out in the *NPGI Five-Year Plan: 2014-2018*

Focus on first year's priorities in database sustainability, data management, and training

- **Identifying strategic research needs** and resource gaps,
- **Prioritizing genomics tools and resources** (including, but not limited to, analytical and genetic tools, sequencing needs, and databases)
- **Defining new strategies** that will meet community needs and priorities sustainably
- **Advance biological innovation and breakthrough discovery**
- **Promote coordination** with international agencies and private sector supporting plant genomics



IWGPG Charge

IWGPG established three Task Forces to implement the goals outlined in the *NPGI Five Year Plan: 2014-2018*

– Big Data

- Access to HPC, data, germplasm; systems interoperability and standards
- Long-term database sustainability

– HTP Phenotyping

- Field-based HTP automated phenotyping
- Data analysis pipelines

– Crop Plant Microbiome

- Imaging and metagenomics technologies
- Plant breeding for sustainable agricultural systems

Discussion - Big Data

- What areas of basic research, tools and resources are needed to facilitate interoperability and promote sharing to advance breakthrough discoveries?
- What needs are not currently being met? In this regard, are there opportunities to leverage existing data, tools and infrastructure?
- What areas of research training and skills are required and not currently being met?
- What opportunities do you see for leveraging investments through public:private and international partnerships?

My Contact Information

Diane Jofuku Okamuro

Program Director

Directorate for Biological Sciences

Division of Integrative Organismal Systems

Plant Genome Research Program

Email: dokamuro@nsf.gov

Phone: (703) 292-4508



Where discoveries begin