

The Saga of the Soybean Genome

- Coordination and planning
 - Hold lots of meetings in airport hotels!
- Results of coordination and planning
 - Get community on board
 - Identify resources and key individuals
 - Establish SoyGEC
- Flow of a generic genome project
- What are we doing in soybean?
- Free suggestions, that should be treated as such.

Coordination & Planning

- Soybean Genomics Workshop
St. Louis, MO, October 21 - 22, 1999 (United Soybean Board-Sponsored)
- Legume Crops Genome Initiative Workshop
Hunt Valley, MD July 30 - 31, 2001 (USDA-Sponsored)
- Strategic Plan for Soybean Genomics Workshop (2003 - 2007) St. Louis, MO
May 20 - 21, 2003 (United Soybean Board-Sponsored)
- Plan for Soybean Genomics Conference
St. Louis, MO October 21, 2003 (National Science Foundation-Sponsored)
- CATG Workshop (Cross-Legume Advances Through Genomics) off-shoot
from the U.S. Legume Crops Genomics Initiative, Santa Fe, NM December
14 - 15, 2004 (USDA- NSF-Sponsored)
- SOYCAP Workshop (Coordinated Agricultural Project)
St. Louis, MO December 16 - 17, 2004 (USDA-Sponsored)

Outcomes of coordination & planning

- Assessment of resources

Genome features

1,115 Mbp/haploid genome
chromosomes: 20/haploid complement
Wide range in chromosome size
Five chromosomes with at least one heterochromatic arm

Special biological features of interest

Ancient paleopolyploid (14 MY and 45-55 MY)
Accumulates high levels of seed protein
Large seed size
Advanced embryogenic systems methodologies
Adapted to wide range of latitudes
Close relatives are one annual (*G. soja*) and
many perennials (*Glycine* sp.)
Allergen-free germplasm exists
Produces many bioactive seed compounds

Outcomes of coordination & planning

- Assessment of resources
- Soybean Genomics Action Plan (USB committee, 2003)

Scientist/Community Role

Establish a scientific Executive Committee
Agree on a standard map and/or standard population(s)
Agree on a 'gold standard' physical map
Agree on standard BAC libraries, and insure easy
accessibility and high quality
Agree on a 'gold standard' genomic database
Immediately organize a genomics workshop involving
the broader plant community

Resource Development Role

Physical and transcript map
Gene-specific arrays
Organize a functional-genomics initiative

Outcomes of coordination & planning

- Assessment of resources
- Soybean Genetics Executive Committee
- Soybean Genetics Executive Committee

Membership and Rotation Schedule:

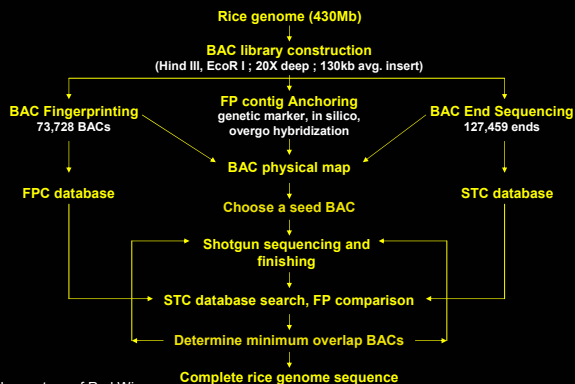
Randy Nelson, USDA, Illinois, 2007
 Jim Specht, Nebraska, 2007
 Randy Shoemaker, Chair, USDA, Iowa, 2006
 Perry Cregan, USDA, Maryland, 2005
 Roger Boerma, Georgia, 2005

Wayne Parrott, Georgia, 2004
 Jim Orf, Minnesota, 2004

The usual flow of a genome project

- Genetic maps
- EST project
- Physical map
- Preliminary genome sequencing
- Map integration
- Determination of sequencing approach
- Sequence genome
- Hope someone will use it.

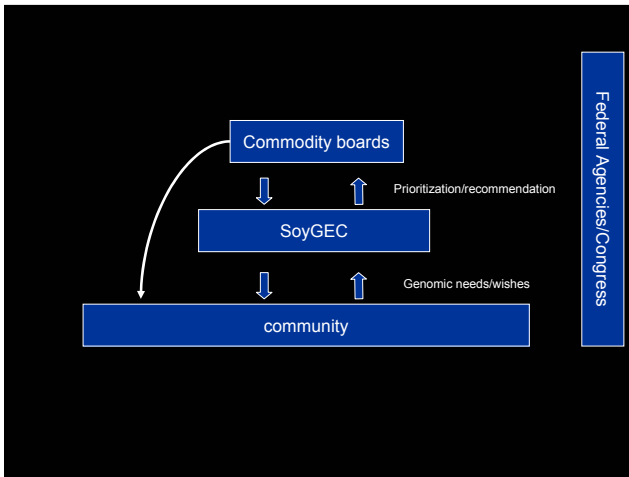
Strategy for sequencing the rice genome



Slide courtesy of Rod Wing

Meanwhile, in soybean...

- Genetic Maps:
 - Funded by commodity boards since mid '80s (lots of sequence-based markers)
- EST project:
 - >300,000 ESTs funded by NSF and USB
- Physical map (NSF & USB):
 - Round 1: wrong genotype
 - Round 2: in progress
- Integration of genetic and physical maps (in progress)
- Preliminary sequencing (Genome survey)
 - GSS seqs >30K (USDA, USB and NSF)
 - Several BACs (industry and community)



Important notes

- Decide on genotype/populations
- Keep community involved
 - Solve problems early!
 - Open lines of communication
- Use experts outside the community
 - This is not the first genome project
- Funding issues
 - Commodity boards were imperative
 - What is the larger biological question (comparative, evolutionary, specific traits?)
- Informatics
- Think beyond just the sequence: it should be a means to an end.
- Who are the end-users? Why do they need a sequence? What type of sequence?