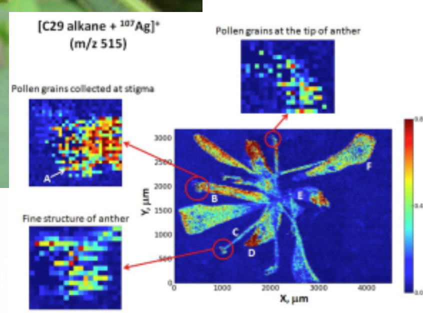
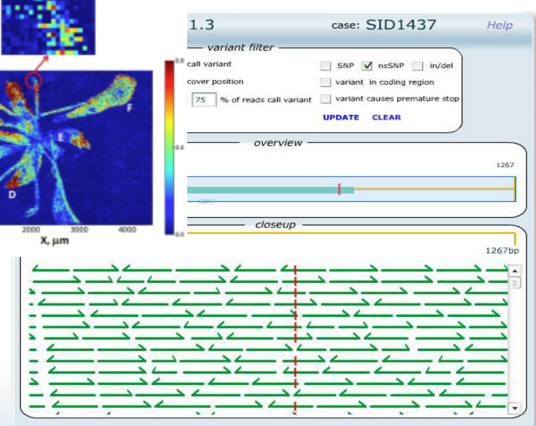
At the Interface of Engineering and Plant Science

International Workshop on Engineered Plants

29 April 2014





$$a_{i,2*j-1} *. L_j = x_{i,1}$$

$$a_{i,2*j} * L_j. = x_{i,2}$$

$$\sum_{j=1}^{10} L_j = 1 \quad \forall_j = 1,2,3 \ldots 10$$

$$\sum_{j=1}^{10} U_j = 1 \quad \forall_j = 1,2,3 \ldots 10$$

Bill Beavis
Lizhi Wang

IOWA STATE UNIVERSITY
Department of Agronomy
Crop, Soil, and Environmental Sciences

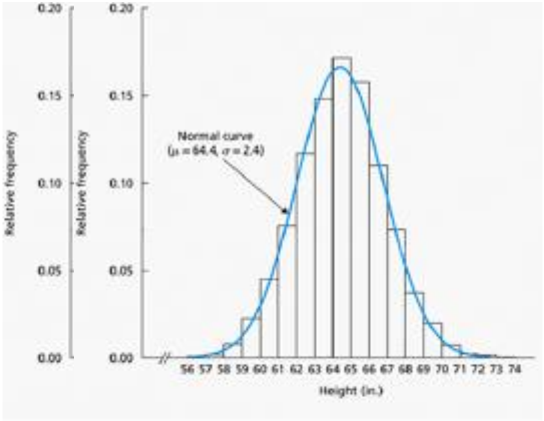
Plant Breeding Systems Engineering

- Plant Breeding at a glance
- Systems Engineering at a glance
- Educational lessons at the interface

Plant Breeding (ΔG): at a Glance

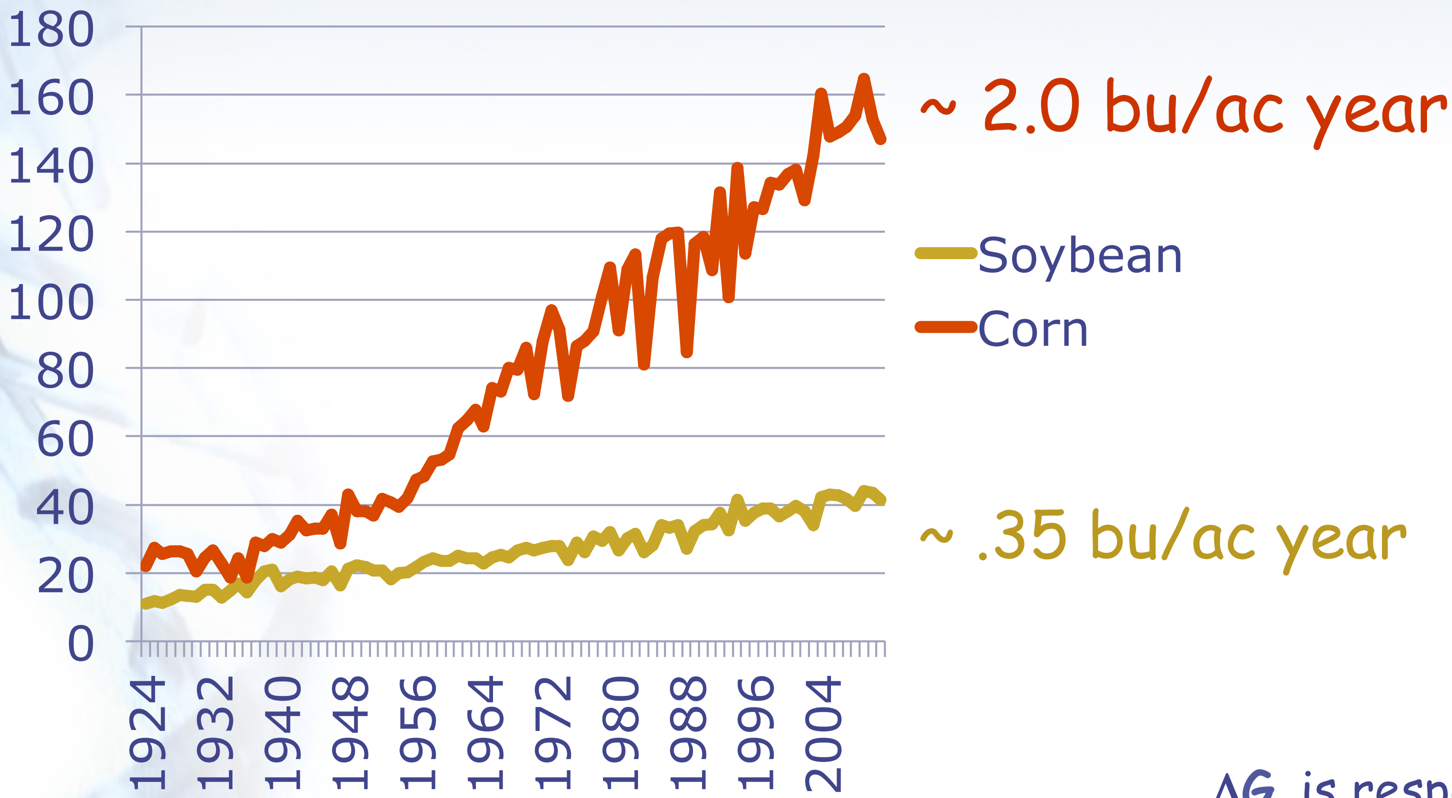
Create Useful Genetic Variability

$$\Delta G = \iota \, r_{RS} \, \sigma_A$$



Select

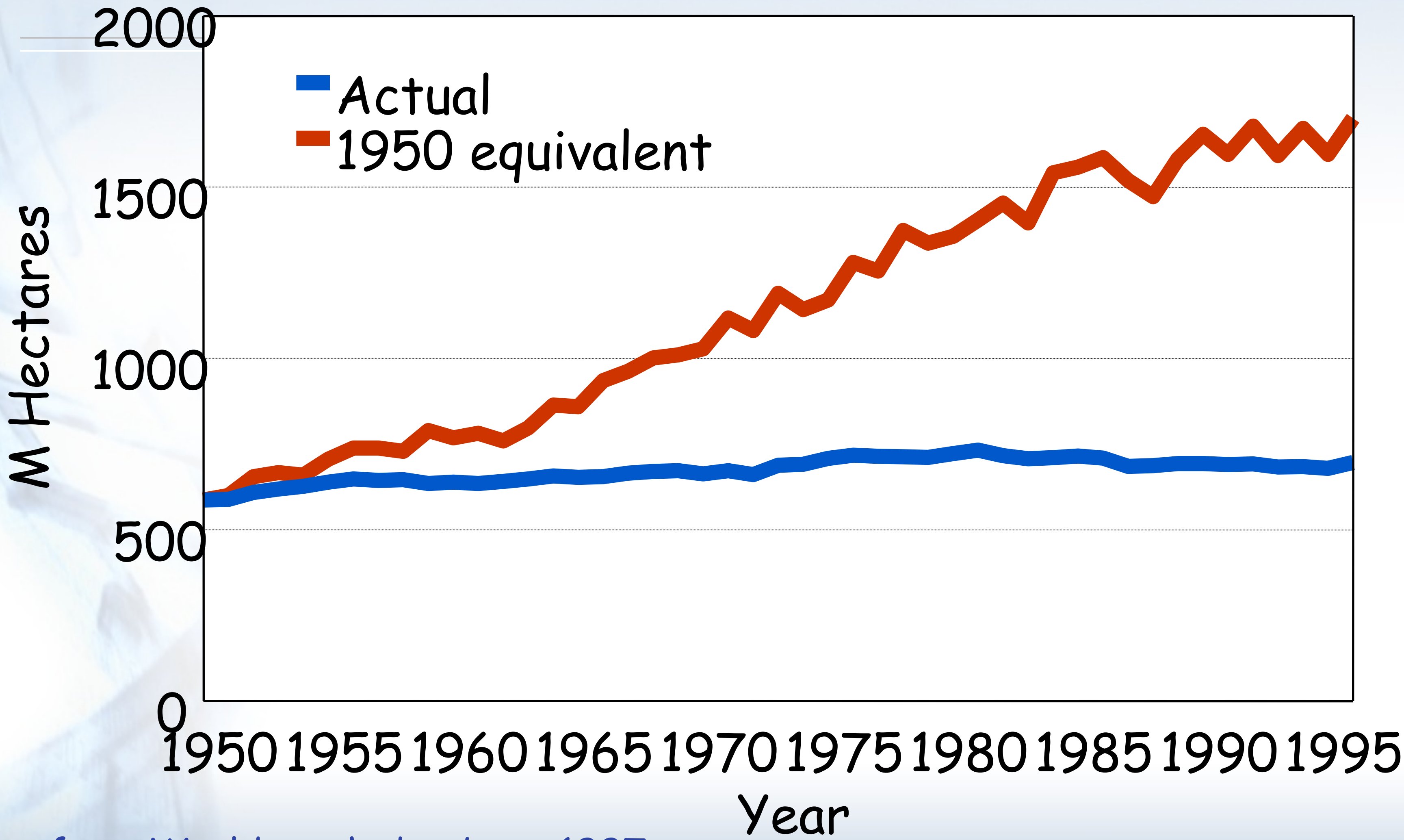
Realized Gains in Corn and Soybean Production



ΔG_c is responsible for up to 80% of the increase -Duvick, 1993, 2005

Provided by Brian Diers, UIUC

Hectares of Grain
With and Without Genetic Improvements



Data from Worldwatch database 1997,
Assembled by Chris Somerville

Lessons:

1. Plant Breeders believe that all food, fuel and fiber production challenges are fundamentally genetics challenges.
2. Given sufficient resources we can discover the underlying genetics needed to meet all food, fuel and fiber production challenges.

Systems Engineering:

Problem Challenge

Model

Algorithm

Implement Solve

Lessons:

3. Systems Engineers believe that all food, fuel and fiber production challenges are fundamentally technical challenges.
4. Given sufficient resources we can design, develop and deliver technologies to meet all food, fuel and fiber production challenges.

Plant Breeders and Systems Engineers
are supremely confident in our
respective approaches and abilities

Crucial Conversations:

Problem Challenge

Parental Prediction (Selection)

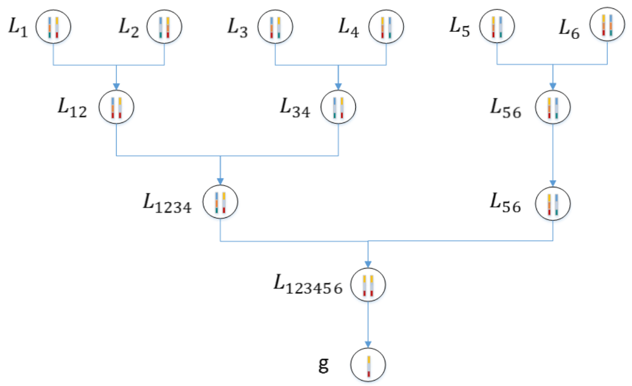
- Consider a finite breeding population
- Each member of the breeding population has a set of diploid chromosomes consisting genetic loci
- At each locus there is a desirable and undesirable allele
- What is the fastest, cheapest breeding strategy to maximize the probability of obtaining a genome consisting of all desirable alleles?
- Predict Value of Parents to produce progeny with the desirable genome

Model

Define Predicted Parental Value

Assign each individual's diploid genotypes to a matrix, x, and it's gamete to g

$$x = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,2k} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,2k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n-1,1} & x_{n-1,2} & \cdots & x_{n-1,2k} \\ x_{n,1} & x_{n,2} & \cdots & x_{n,2k} \end{bmatrix} \Rightarrow g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_{n-1} \\ g_n \end{bmatrix}$$



$x_{i,j} \in \{0,1\}$ is a random variable indicating the desirability of the allele in locus $\{i,j\}$

$g_i \in \{0,1\}$ is a random variable indicating the desirability of the inherited allele.

Model

Define Predicted Parental Value

$a_{i,2*j-1} * L_j = x_{i,1} \quad \forall_{i=1,2..10}$

$a_{i,2*j} * L_j = x_{i,2} \quad \forall_{i=1,2..10}$

$a_{i,2*j-1} * U_j = x_{i,3} \quad \forall_{i=1,2..10}$

$a_{i,j} * U_j = x_{i,4} \quad \forall_{i=1,2..10}$

$$\sum_{j=1}^{10} L_j = 1 \quad \forall_j = 1,2,3 \ldots 10$$

$$\sum_{j=1}^{10} U_j = 1 \quad \forall_j = 1,2,3 \ldots 10$$

$w_{1,j} \le 0.25$

$w_{i,j} \le x_{i,j}$

$w_{i,j} \ge 0$

$x_{i,j} \ge 0$

L_j, U_j *binary*

$$w_{i,1} \le (1-r_{i-1})^2 * w_{i-1,1} + r_{i-1} * (1-r_{i-1}) * w_{i-1,2} + 0.5 * r_{i-1} * w_{i-1,3} + 0.5 * r_{i-1} * w_{i-1,4} \quad \forall_i = 2,3,4 \ldots 10$$

$$w_{i,2} \le r_{i-1} * (1-r_{i-1}) * w_{i-1,1} + (1-r_{i-1})^2 * w_{i-1,2} + 0.5 * r_{i-1} * w_{i-1,3} + 0.5 * r_{i-1} * w_{i-1,4} \quad \forall_i = 2,3,4 \ldots 10$$

$$w_{i,3} \le 0.5 * r_{i-1} * w_{i-1,1} + 0.5 * r_{i-1} * w_{i-1,2} + r_{i-1} * (1-r_{i-1}) * w_{i-1,3} + (1-r_{i-1})^2 * w_{i-1,4} \quad \forall_i = 2,3,4 \ldots 10$$

$$w_{i,4} \le 0.5 * r_{i-1} * w_{i-1,1} + 0.5 * r_{i-1} * w_{i-1,2} + (1-r_{i-1})^2 * w_{i-1,3} + (1-r_{i-1})^2 * w_{i-1,4} \quad \forall_i = 2,3,4 \ldots 10$$

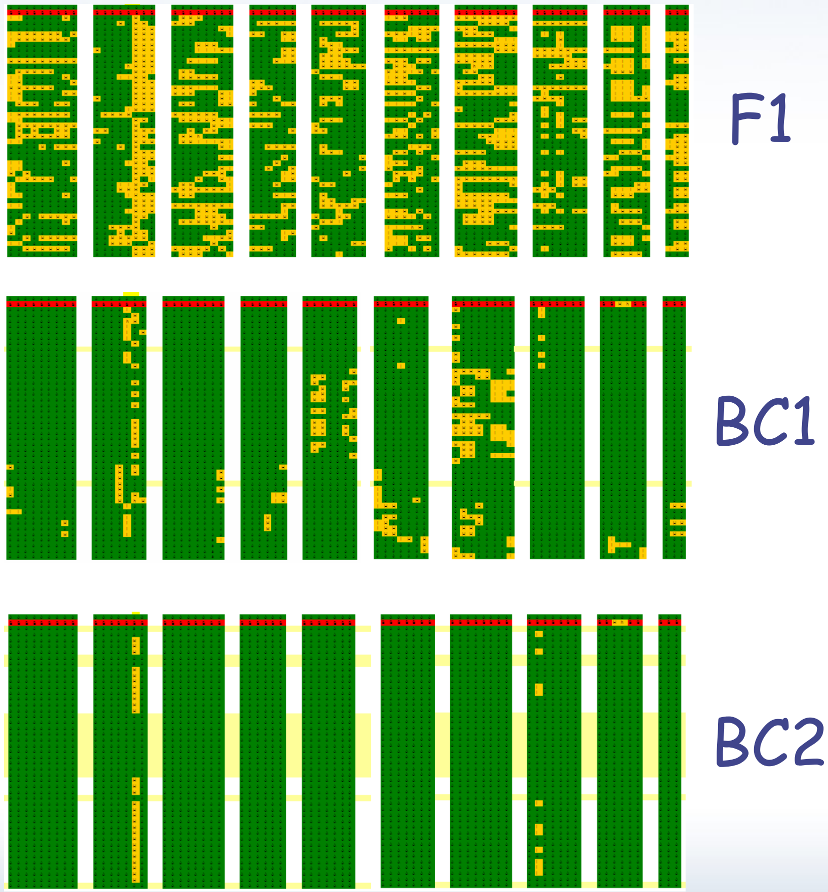
$$PPV = Max\left(\sum_{i=1}^{10} \sum_{j=1}^{4} w_{i,j}\right)$$

Possible breeding strategies

- Recurrent Genomic Selection,
  - Genomic Prediction
- Multi-phase
  - Gene stacking using Backcrossing and F2 enrichment
- Dynamic Decision Process
  - Combine desirable alleles from 5 donors based on genome-wide marker assays and Predicted Parental Breeding Values.

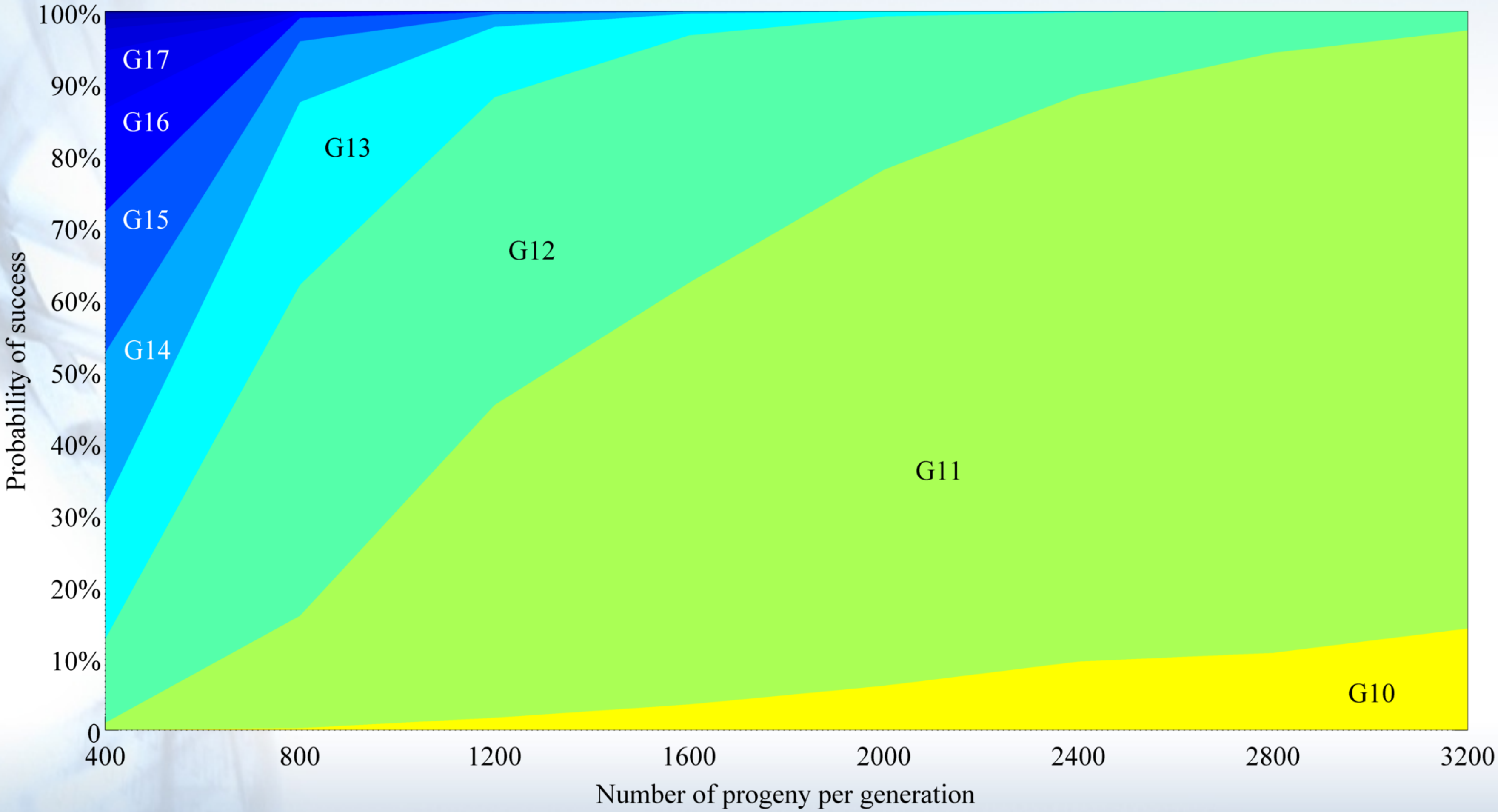
Algorithm

Gene Stacking with Backcrossing and F2 Enrichment



Implement
Solve

Marker Assisted Backcrossing And F2 Enrichment

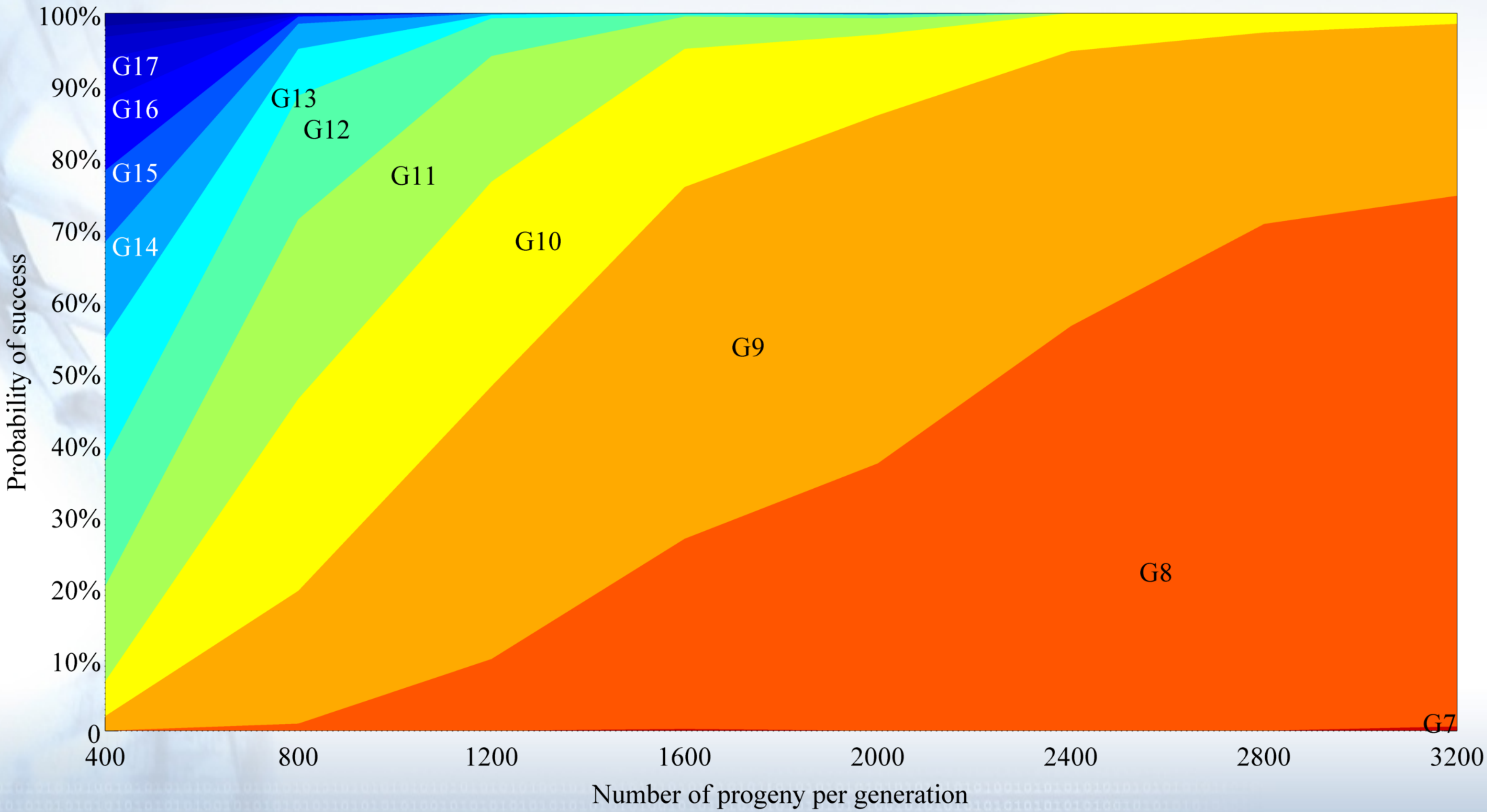


Algorithm

Dynamic Decision Process Using Predicted Parental Values

Dynamic Decision Process Using Predicted Parental Values

Implement
Solve



Lessons:

5. Systems engineering students are capable of explaining objective functions to plant breeders.
6. Plant breeding students are capable of explaining genetics to systems engineers.
7. Content is important, but pedagogy based on appropriate Applied Learning Activities (ALA's) provides motivation for Biologists to learn Math.

Lessons:

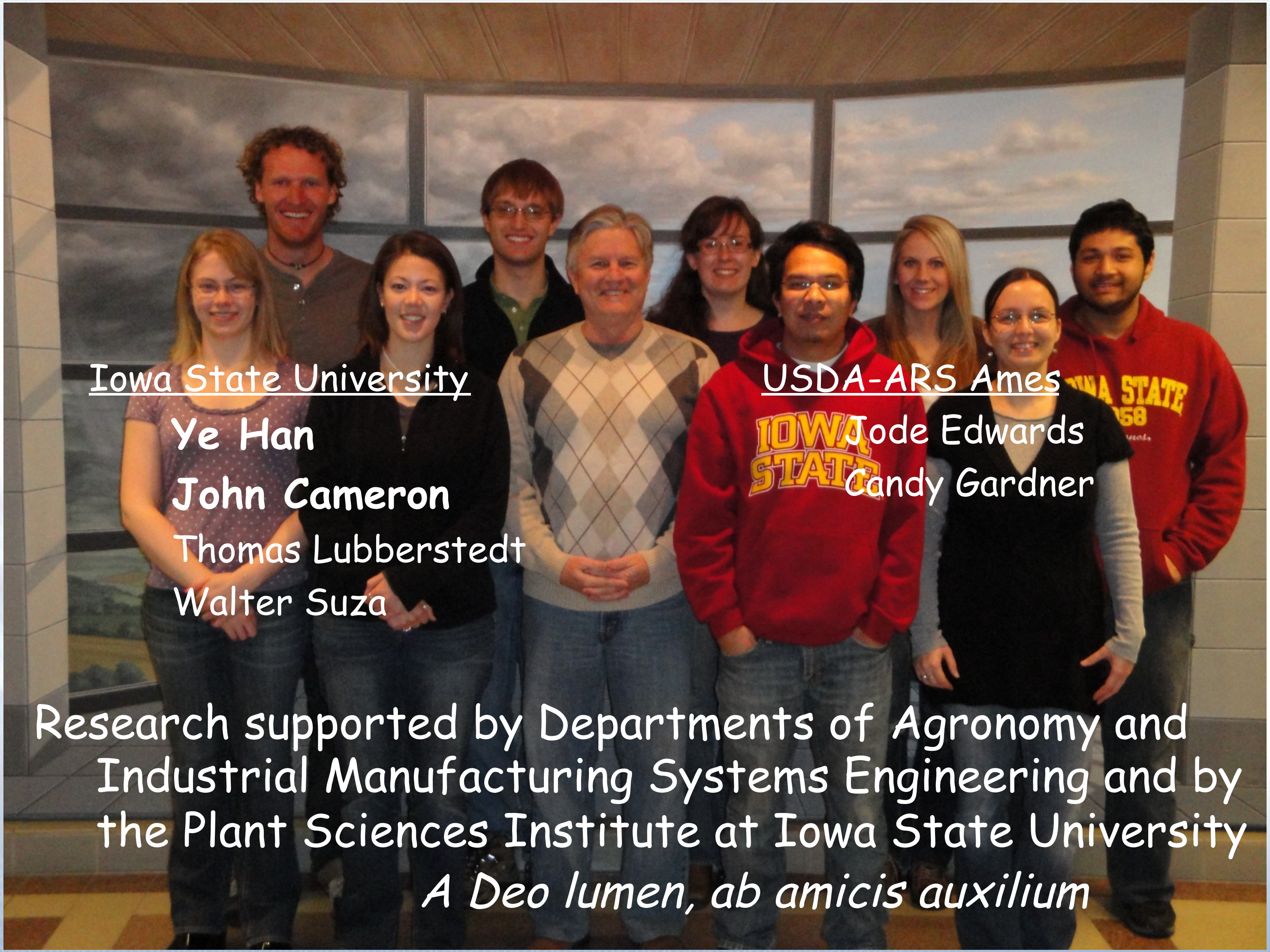
8. Language of Systems Engineers is precise and often supported by mathematical proofs
9. Language of Plant Scientists is based on loose definitions supported by context.
   - (Genotype? It depends)
   - (Ontology for phenotypes? Maybe)
   - Biology is in the midst of a Kuhnian Revolution: what is a gene?
10. While frustrating for those who need established foundations, now is the time to make significant scientific contributions.

An Emerging Opportunity:

Theory for bi-level optimization:
- Two propositions.

Highly relevant for:
- Hierarchical evolutionary dynamics
- Engineering metabolic flux
- Breeding for adaptation to rapidly changing environments

Acknowledgements



Iowa State University

Ye Han

John Cameron

Thomas Lubberstedt

Walter Suza

USDA-ARS Ames

Jode Edwards

Candy Gardner

Research supported by Departments of Agronomy and Industrial Manufacturing Systems Engineering and by the Plant Sciences Institute at Iowa State University

A Deo lumen, ab amicis auxilium