Why breeding? Need to increase yield/acre

Plant Breeder – develops products for farmers and industry to meet stakeholder requirements and basic human needs

- **Yield** can be **increased** *per se*
- **Yield losses** can be **reduced** (biotic and abiotic factors)

What is important -

- **net farm income**, sustainable and **environmentally friendly**
1/3 bushels/acre/yr yield increase

\[ S_i = \text{incident solar radiation during a growing season} \]  
\[ (0.487 \text{ is the fraction available for photosynthesis}), \]
\[ \varepsilon_i = \text{radiation interception efficiency of the crop canopy} \]  
\[ (\text{speed and duration of canopy closure along with canopy size and architecture}), \]
\[ \varepsilon_c = \text{efficiency of converting intercepted radiation into biomass}, \]
\[ \varepsilon_p = \text{efficiency of partitioning biomass into a harvestable product or the harvest index}. \]

In the US, yield improvement in major soybean growing regions has come from

- lowering protein (increasing \( \varepsilon_c \)),
- later maturity (increasing \( \varepsilon_i \)),
- reduced height (increasing \( \varepsilon_p \))
- stronger stem (increasing \( \varepsilon_i \) and \( \varepsilon_p \)),
- higher HI (\( \varepsilon_p \))
..but working with an immobile entity (plants) and a variable environment.

- Soil
- Air
- Pathogen and symbionts (above or below ground)
  - Existing pathogens
  - New and emerging pathogens

Environment factors are becoming more variable:
- Temperature (spatial and temporal)
- Precipitation (rain and snow) (spatial and temporal)
- Growing days
The colors on the map show temperature changes over the past 22 years (1991-2012) compared to the 1901-1960 average, and compared to the 1951-1980 average for Alaska and Hawai‘i. The bars on the graphs show the average temperature changes by decade for 1901-2012 for each region.

The colors on map show annual total precipitation changes for 1991-2012 compared to 1901-1960 average. The bars on the graphs show average precipitation differences by decade for 1901-2012 for each region. The far right bar in each graph is for 2001-2012.

The frost-free season length, defined as the period between the last occurrence of 32°F in the spring and the first occurrence of 32°F in the fall, has increased in each U.S. region during 1991-2012 relative to 1901-1960.

Increases in frost-free season length correspond to similar increases in growing season length.

Why these variable parameters important for yield protection (plant breeding)?

• In 2014 alone, soybean yield loss estimated due to pathogen and pest was $3.9 billion (http://extension.cropsciences.illinois.edu/fielddcrops/diseases/yield_reductions.php; assuming $8 per bushel price)

• Pathogens don’t sleep
  – adapt, virulent strains emerge and predominate
  – previously identified R genes may become ineffective
Single or few trait targeted breeding: leaving yield on the ground

<table>
<thead>
<tr>
<th>Trait</th>
<th>If 2% yield loss; $ Loss per acre</th>
<th>If 5% yield loss; $ Loss per acre</th>
<th>Estimated yield loss 2500 acres (@5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average yield potential = 45 bushels/ac; average price =$10/bushel)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCN</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td>SDS</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td>Phytophthora rot</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td>BSR</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td>WM</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td>IDC</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td>Soybean Aphid</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td>Viruses</td>
<td>$9/acre</td>
<td>$22.5/acre</td>
<td>-($56,250)</td>
</tr>
<tr>
<td><strong>If more than one</strong></td>
<td><strong>-[(($9/acre)*# of traits]</strong></td>
<td><strong>-[(($22.5/acre)*# of traits]</strong></td>
<td><strong>?????</strong></td>
</tr>
</tbody>
</table>
Overview of Plant Breeding Process

• Set objectives (How?)
• Assemble genetic variation to meet the objectives (From where? New tools to increase the “good” diversity)
• Generation advancement, selection (New tools to increase the effectiveness of selection?)
• Breeder seed production
• See it through commercialization
SET OBJECTIVES, IDENTIFY PARENTS
DECIDE ON HYBRIDIZATIONS (ON OBJECTIVES; COMPLEMENTING)
GENERATION MANAGEMENT

Training set
(use genotype and phenotype)

Develop Prediction Model

Validation set
(use genotyped only)

Make predictions and selections in the breeding program
Update the training population

IDENTIFY SUPERIOR CULTIVAR; BREEDER SEED; COMMERCIALIZATION

MAS, WN, GWS, HTP

FORWARD BREEDING EXAMPLE

Soybean breeding 101

IOWA STATE UNIVERSITY
Department of Agronomy

i’m an agronomist
applying science to fuel & feed our global society
- Set objectives, identify parents
- Decide on hybridization (complementing objectives)
- Generation advancement
- Selections made: single plants, rows or small plots (phenotypic and/or genotypic – MAS / GWS)
- Genetic variability decreases / generation

Reduced population size
Replicated trials, multi-env

Identify superior cultivar; breeder seed; commercialization

Mas, WN, GWS, HTP
Setting objectives

- Setting objectives allows the breeder to make strategic decisions such as:
  - Parents that have the necessary complementation of traits to develop progeny that possess desirable traits from both parents
  - Breeding method
  - Selection strategy and plan for any specialized nursery or tools
  - Traits and generation of selection for each
<table>
<thead>
<tr>
<th>Disease / pest resistance</th>
<th>Breeding and research priority 1</th>
<th>Breeding and research priority 2</th>
<th>Breeding and research priority 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCN</td>
<td>Charcoal rot</td>
<td>Phomopsis</td>
<td></td>
</tr>
<tr>
<td>SDS</td>
<td>Bacterial Blight</td>
<td>Stem canker</td>
<td></td>
</tr>
<tr>
<td>Phytophthora stem/root rot</td>
<td>Frogeye</td>
<td>Alfalfa mosaic virus</td>
<td></td>
</tr>
<tr>
<td>WM</td>
<td>Root/seedling rot (Fusarium, Phythium, Rhizoctonia)</td>
<td>Cercospora blight</td>
<td></td>
</tr>
<tr>
<td>Brown Stem rot</td>
<td>Downey mildew</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bean pod mottle virus</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soybean mosaic virus</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soybean Vein necrosis virus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insect resistance</td>
<td>Soybean Aphid</td>
<td>Japanese beetle</td>
<td>Kudzu bug</td>
</tr>
<tr>
<td></td>
<td>Brown marmorated stink bug</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thrips</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bean leaf beetle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abiotic stress</td>
<td>Water stress</td>
<td>Micro-organism</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Temp. stress</td>
<td>IDC</td>
<td></td>
</tr>
<tr>
<td>Quality and nutrition</td>
<td>Oil and Fatty acid profile</td>
<td>Carbohydrate</td>
<td>Phytate</td>
</tr>
<tr>
<td></td>
<td>Protein</td>
<td></td>
<td>Nutritional profile</td>
</tr>
<tr>
<td>Genomics</td>
<td>Genomic selection</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marker assisted targeted breeding</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gene stacking</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genetic variability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment</td>
<td>Genotype x Environment</td>
<td>Agronomic practices</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prediction (Site)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Off-season nursery</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Residue traits</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Form the genetic base

- Advanced lines from the breeding program
- Advanced lines from another breeding program
- Released cultivar
- Germplasm line from gene bank or a pre-breeding program
- Introductions (from other countries) from colleagues or gene bank
- Mutant line populations (unselected or selected)
- Wild relative (need to be crossable or resources available to do embryo rescue if needed)
Genetic variation in soybean

- 80 (<0.02%) of ~45,000 landraces account for 99% of the collective parentage of North American soybean cultivars released between 1947 and 1988 (Carter et al. 2004).
- 17 of these 80 account for 86% of the collective parentage, with the remaining 63 landraces contributing <1% each (Gizlice et al. 1994).
- What should be our approach to bring genetic diversity? “good” vs “bad” diversity?
Single plant selection

Maximum allowable testing to assess adaptation and genetic worth

Maximum genetic variation

New cultivar (1-few)
Generations and cultivar development: wheat and soybean?

Yield, Quality

Disease resistance:
- Leaf Rust
- Stem Rust
- Common Bunt
- Fusarium Head Blight
- Leaf spots
- Loose Smut
- Stripe rust
- Ergot
- Smudge/blackpoint

Insect Resistance:
- Midge
- Wheat stem Sawfly
- Hessian Fly

Yield, Quality, Adaptation

Disease resistance:
- Sudden death syndrome (SDS)
- Brown stem rot (BSR)
- Phytophthora stem rot
- Phytophthora root rot

Insect and Pest Resistance:
- Soybean cyst nematode (SCN)
- Soybean aphid

Iron deficiency chlorosis (IDC)
- Water stress
Advancing generations and selection

- Choose an appropriate breeding method to develop inbreeding populations.
- All methods in pureline breeding lead to an increase in homozygosity, a reduction in the genetic variance within families, and an increase in the genetic variance between families (genetic variability exists among, but not within lines).
- Populations will be ultimately composed of an array of different inbred homozygous lines (pure lines).
- Cultivar development is aimed at identifying the best purelines.

- Factors to consider: starting generation to select, use of markers or phenotype to enrich desirable alleles, DISCARD or SELECT?, fewer locations are used in early generations (Why?) so carefully choose the locations (how?)
- Response to selection: What strategy to use?
- Experimental designs?
Large scale yield testing, seed multiplication

• Final stages in the breeding cycle will involve lines that are considered pure lines (non-segregating). At this stage, more extensive testing of the few best recombinants from a cross is done for agronomic performance and end-use quality or other complex traits.

• Product placement information is obtained. Multi-environment testing is done for adaptation and stability, and environments may be locations or a combination of locations and years.
New tools in phenotyping
High throughput phenotyping (HTP)

- **Plant phenotyping:** application of methodologies and protocols to measure specific traits, ranging from the cellular to whole plant or canopy level, related to plant structure and function.

- The bottleneck in phenotyping has led to HTP approaches, unlocking new prospects for non-destructive field-based phenotyping.

- Breeding/Res. + Precision Ag.
Breeding programs phenotype large numbers of plants in each crop cycle and these measurements have a requirement of time sensitivity and are growth stage dependent.

Aerial, ground? Which traits? What sensors?
HTP Examples

Imaging sensors:
• Digital images
• Multi-spectral
• Hyperspectral images

• Manned ground system
• Unmanned ground system
Post processing steps:

• image alignment and segmentation,
• Geometric, radiometric, atmospheric correction (corrections related to the coordinates, sensitivity of the sensor or camera, topography and sun angle, and atmospheric scattering and absorption),
• mosaicking,
• algorithms for trait extraction
Remote sensing phenotyping methods are non-destructive, non-invasive approaches based mostly on the information provided by visible/near-infrared (VIS-NIR) radiation reflected (or transmitted) and far-infrared (thermal) radiation emitted by the crop.
**RGB cameras:** Red, green, and blue light (called visible or RGB) imagery enables the estimation of green biomass (NDVI type of information). Trait extraction of phenotype that can generally be visibly distinguished.

**Multispectral cameras:** for crop monitoring via remote sensing. Limited number of spectral bands in the VIS-NIR regions. Can be used for vegetation indices, senescence evaluation, nutrient status, water content.

**Hyperspectral camera:** acquisition of hundreds of images at once, covering 300 to 2500 nm. Very extensive trait generation, yield and disease signatures.

**Long-wave infrared cameras or thermal imaging cameras:** thermal imaging in phenotyping includes predicting water stress in crops, disease detection in plants.

I think my Nest smoke alarm is going off. Google AdWords just pitched me a fire extinguisher and an offer for temporary housing.
Machine learning and Deep Learning

• “Machine learning (ML) is a method of data analysis that automates analytical model building. Using algorithms that iteratively learn from data, machine learning allows computers to find hidden insights without being explicitly programmed where to look.”

• In ML, computers apply statistical learning techniques to automatically identify patterns in data (for ex, images), and these techniques can be used to make highly accurate predictions. [ML is about learning from the past data to do better in the future. Pattern or feature identification (application in imaging analysis in plant science)].

• “Deep learning is a fast-growing area in ML research that has achieved breakthroughs in speech, text and image recognition. It enabling a computer to learn tasks, organize information and find patterns on its own.”

ICQP solution using machine learning for plant stress phenotyping
HTP-ML applications

- Identify useful germplasm
- Identify genetic factors
- Time series enabled accurate phenotyping
- Simultaneous phenotyping at multiple locations
- Decision making for parents to pick for hybridization, selection of improved genetics
Needs of a plant breeding program

- Mandate, Objectives, Genetics
- People
  - Field and lab
- Field and lab equipment
  - Small plot equipment
  - NIRS, HPLC, GC
  - Molecular lab
- Infrastructure (space, storage, testing sites etc)
- New way of thinking (optimization, big data, HTP, ML approaches, GWS...)
New trends....

• Gaining importance of Prediction science in plant breeding (combine genetics, phenomics, environment, crop modelling)
• Better engineering tools and solutions to address bottlenecks
• All of these (models, predictions) will require high quality phenotyping
• New disciplines and expertise will be a component of plant breeding programs
Funding Support

- Iowa Crop Improvement Association
- Iowa Soybean Association
- Iowa State University
- Monsanto
- North Central Soybean Research Program
- R.F. Baker Center for Plant breeding
- United Soybean Board
You are in the right discipline of career when you feel you haven’t worked a day in your life!

Thank you!!