Cultivar Development

Walter Suza (Editor); Kendall Lamkey (Editor); and Rita H. Mumm

Iowa State University Digital Press
Ames, Iowa

How to cite this publication:

Suza, W., & Lamkey, K. (Eds.). (2023). Cultivar Development. Iowa State University Digital Press. https://doi.org/10.31274/isudp.2023.144

This is a publication of the
Iowa State University Digital Press
701 Morrill Rd, Ames, IA 50011
https://www.iastatedigitalpress.com
digipress@iastate.edu

CONTENTS

About the PBEA Series

Background

The Plant Breeding E-Learning in Africa (PBEA) e-modules were originally developed as part of the Bill & Melinda Gates Foundation Contract No. 24576.

Building on Iowa State University's expertise with online plant breeding education, the PBEA e-modules were developed for use in curricula to train African students in the management of crop breeding programs for public, local, and international organizations. Collaborating with faculty at Makerere University in Uganda, University of KwaZulu-Natal in South Africa, and Kwame Nkrumah University of Science and Technology in Ghana, our team created several e-modules that hone essential capabilities with real-world challenges of cultivar development in Africa using Applied Learning Activities. Our collaboration embraces shared goals, sharing knowledge and building consensus. The pedagogical emphasis on application produces a coursework-intensive MSc program for Africa.

PBEA Project Director: Walter Suza

Original Module Coordinator: Rita H. Mumm

Collaborating Faculty and Experts in Africa: Richard Akromah, Stephen Amoah, Maxwell Asante, Ben Banful, John Derera, Richard Edema, Paul Gibson, Sadik Kassim, Rufaro Madakadze, Settumba Mukasa, Margaret Nabasirye, Daniel Nyadanu, Thomas Odong, Patrick Ongom, Joseph Sarkodie-Addo, Paul Shanahan, Husein Shimelis, Julia Sibiya, Pangirayi Tongoona, Phinehas Tukamuhabwa.

The authors of this textbook series adapted and built upon the PBEA modules to develop a series of textbooks covering individual topic areas. It is our hope that this project will facilitate wider dissemination and reuse of the PBEA modules' content.

Explore the Series

- Crop Genetics
- Quantitative Methods for Plant Breeding
- Molecular Plant Breeding
- Quantitative Genetics for Plant Breeding
- Crop Improvement
- Cultivar Development

Chapter 1: Plant Breeders Rise to the Challenge of Feeding the World

Rita H. Mumm

Cultivars are the result of plant breeding, the science of applying genetic principles to improve plants for human use.

Plant breeding impacts the life of every individual because it involves the creation and manipulation of economically important traits in plants used for food, animal feed, fiber for clothing and wood products, fuel, and landscaping. Furthermore, plant breeding has been enormously successful! Average corn grain yields in the USA have increased from ~ 1.8 t/ha (~26.8 buA) in the 1930s when hybrids replaced open-pollinated varieties to ~11.7 t/ha (~174.6 buA) today. That is more than a 6-fold increase! The uniformity associated with hybrids is shown in Fig. 1.

Fig. 1 Maize grain production with an elite hybrid. Photo by Anthony Assibi Mahama, courtesy of Iowa State University.

Global Grain Yield Increases

Globally, the average rates of grain yield increase per year for maize, rice, wheat, and soybean are 1.6%, 1.0%, 0.9%, and 1.3%, respectively. Observed global yields from 1961-2021 are shown in Fig. 2. Steady increases are evident with each crop, yet these rates of increase are not sufficient to meet the demand anticipated by 2050.

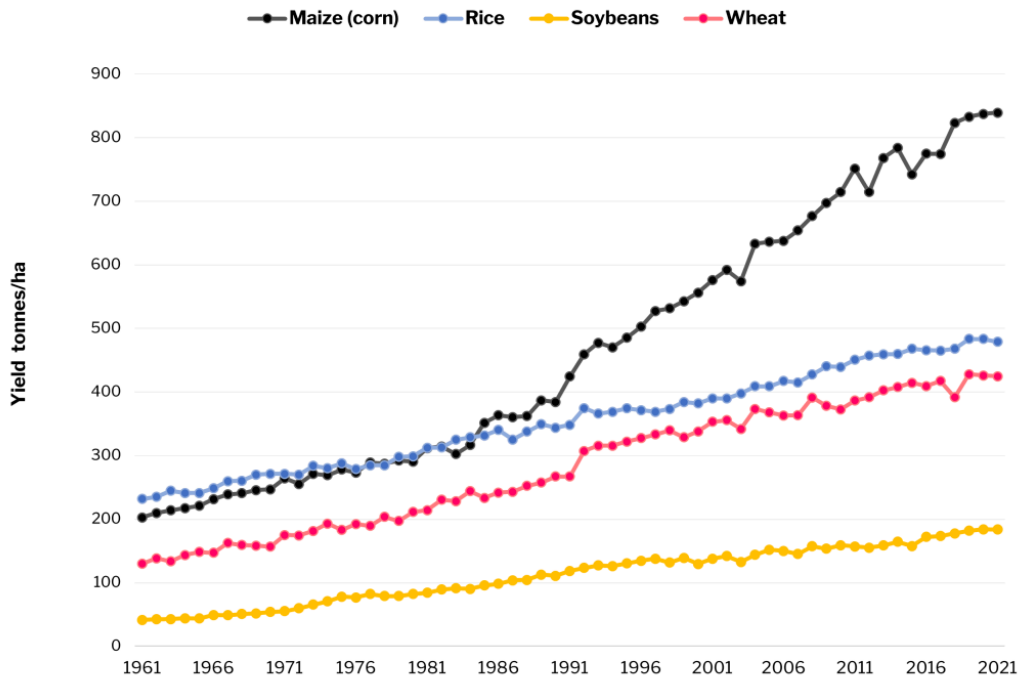


Fig 2a. Global observed crop yields. Data from FAOSTAT.

Grain Yield Increases to Meet Future Global Demand

A 2.4% per year rate of yield gain is needed across crops. The shaded area in Fig 2b shows the trend of the 2.4% yield improvement required each year to meet demand anticipated in 2050, without bringing additional land under cultivation, starting in the base year of 2021.

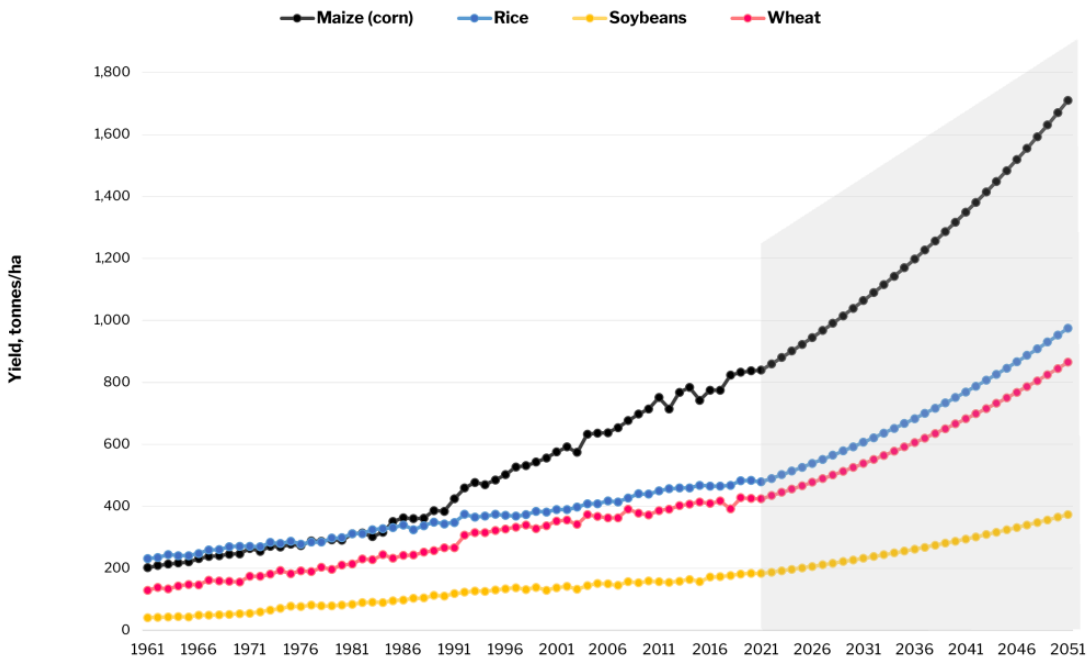


Fig 2b. Global observed crop yields. Data from FAOSTAT.

The Grand Challenge

The Grand Challenge that drives the development of improved cultivars is many-faceted:

- The world population is estimated at more than nine billion by 2050.
- Increased meat consumption in emerging economies as the standard of living increases.
- No appreciable change in available cropland globally and much of the available land is being degraded.
- Falling water tables globally.
- Climate change increases risk in crop production.

Crop yields must increase twofold by 2050 to meet the projected global demand for food and feed (relative to the base year 2008).

Maximizing Agricultural Production

What are some of the ways to achieve greater food production? How can agricultural productivity be improved? More land can be brought into cultivation (although this is not always feasible). Alternatively, more can be produced from each unit of land.

How is this accomplished?

- Improve the genetics of the seed.

- Better production practices to provide the sunlight, water, and soil nutrients plants need, and to mitigate stress factors.

Agricultural production can be maximized when the crop's **yield potential** is manifested.

The Yield Gap

The **yield gap** considers the difference between:

- **Yield potential**, which is the yield productivity potential of an adapted cultivar when grown under favorable conditions without growth limitations from water, nutrients, pests, disease, and other stress factors, and
- **Current realized yield**, which is the actual yield on a specified spatial and temporal scale.

For more information on crop yield gaps, see Lobell, Cassman, & Field (2009).

A Glimpse of Yield Potential

The yield potential of a cultivar could also be considered in light of the biological productivity limit for its crop species, although plant breeders have not hit a permanent ceiling in any crop as yet!

U.S. farmer David Hula produced the highest corn grain yield on record, making him first place winner in the 2015 National Corn Growers Association contest. The 532.0271-bushel-per-acre (35.78 t/ha), certified yield on a 10-acre field was produced in Charles City, Virginia, under reduced tillage (no-till/strip-till) and irrigation with Pioneer hybrid P1197AM planted at a population density of 57,000 seeds per acre (54,500 plants per acre).

Genetic Challenges

Other challenges in plant breeding arise because most key traits of interest (e.g. yield) are **polygenic**. That is, typically many genes are involved in the expression of economically important traits.

- Each gene is thought to contribute a small effect.
- Genetic effects can be difficult to measure due to environmental noise.
- The expression of some genes is influenced by the environment.
- Genes of parents are randomly shuffled when a cross is made!

Undernutrition and Malnutrition

In addition to **undernutrition** caused by not having enough food, over one billion people in the world suffer from **malnutrition** (i.e. "hidden hunger"). Africa is particularly vulnerable. And, among the undernourished, children are especially hard-hit.

Status of Undernutrition and Malnutrition

Facts about undernutrition and malnutrition:

- 3.5 million maternal and child deaths could be prevented annually with improved nutrition.
- In developing countries, iron deficiency affects half of children under age 5, impairing growth, cognitive development, and immune function.
- Vitamin A deficiency affects at least 100 million children, limiting their growth, weakening their immunity, and in acute cases, leading to blindness.
- More than one-third of all African children suffer stunting (low height for weight, irreversible after age 2,) due to malnutrition and undernutrition.
- Stunting in early life is associated with lifetime debilitating neurological effects such as poor cognition and learning, low adult wages, lost productivity, and increased risk of chronic disease.
- Undernutrition, especially during the critical window from conception to 2 years of age, is associated with lower human capital.
- What is more, the devastating effects of malnutrition and undernutrition are across generations: a girl who was fed poorly as an infant is likely to have offspring with lower birth weight.

Crop improvement must be directed to producing *better* food as well as *more* food.

The Cycle of Cultivar Improvement

Generally, cultivar improvement involves the creation or assembly of useful genetic diversity and ways to exploit this variation to achieve targeted breeding goals.

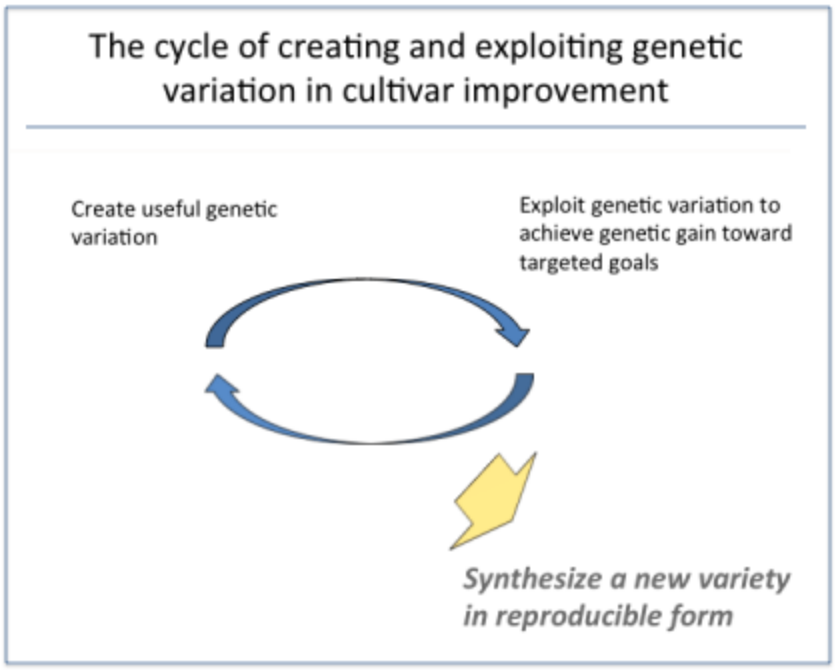


Fig. 3 The cycle of creating and exploiting genetic variation in cultivar improvement. Adapted from Mumm, 2013.

The general philosophy is simple: cross the "best" parents, produce progeny, and then identify and recover progeny that surpass the parents and demonstrate outstanding performance.

The superior progeny may become the basis of a new, improved cultivar; seed volumes (or plant propagules in the case of clonal crops) are produced for distribution.

Typically, this cycle is repeated multiple times to achieve a particular breeding target. Superior progeny may be used as parents in the next cycle, accumulating gains from selection.

Cultivar Improvement Challenges

Challenges include:

- Choosing the "best" parents.
- Identifying the truly superior progeny.
- Environmental noise that reduces **heritability** and makes it difficult to discern performance differences.
- Effective or efficient screens to measure performance for certain traits.
- A lack of knowledge about the **metabolics** and genetic architecture underlying the traits of interest.

Tools for Cultivar Development

Tools help!

For example, prediction can aid in choosing parents; **transformation** and **gene editing** may be used to create new useful genetic variation; doubled haploidy speeds development of homozygous progeny.

DNA-based technologies like molecular markers and sequence information enable genomics-assisted selection; analytics involves automated high-throughput analysis (think grain composition analysis); **phenomics** facilitates evaluation of performance in specialized testing environments (often focused on developing stress-tolerant cultivars).

Knowledge of Genetic Architecture Guides the Plant Breeder

With DNA-based information, knowledge related to genetic architecture and genome function then becomes a part of the cycle, leading to the creation of useful genetic variation and ways to exploit it in cultivar improvement (Fig. 4).

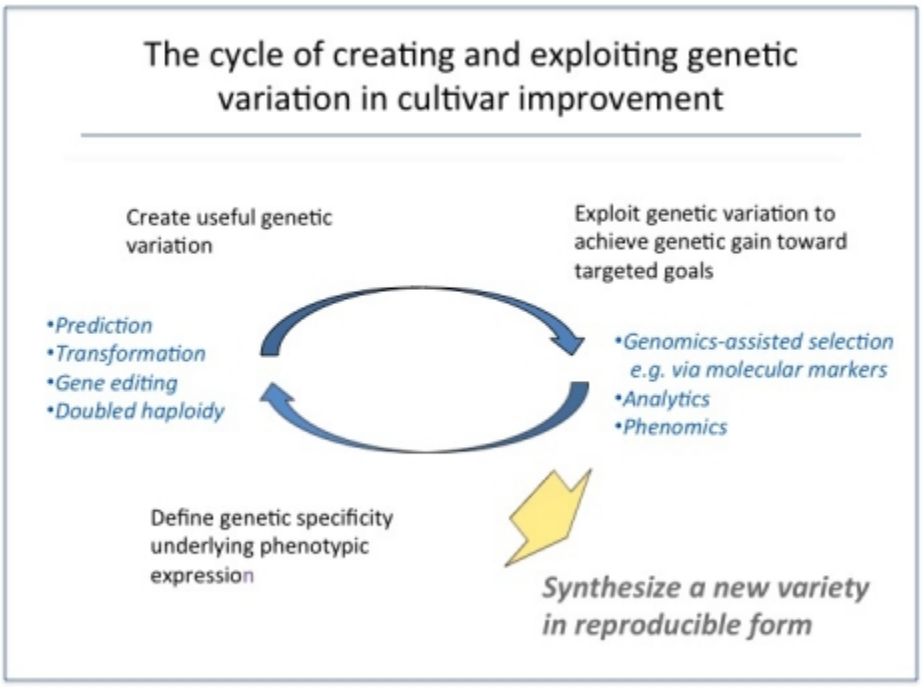


Fig. 4 Inclusion of modern tools in the cycle of creating and exploiting genetic variation in cultivar improvement. Adapted from Mumm, 2013.

Increasing the Frequency of Favorable Alleles

Plant breeders (Fig. 5) are on a mission to increase the frequency of favorable alleles…

and to decrease the frequency of unfavorable alleles!

Fig. 5 CIMMYT scientists develop improved maize cultivars. Photo courtesy of CIMMYT.

Fitness and Gene Frequency

In nature, there is natural selection for “fitness” contributing to the next generation (Fig. 6).

Fig. 6 Gene frequency is skewed in favor of genetics from the fittest individuals. Photo courtesy of USDA.

Selection and Selection Response

In deploying the cycle of creating and exploiting genetic variation to develop improved cultivars, the plant breeder implements basic genetic principles.

- Top-performing individuals from base population are selected as parents to produce the next generation, and that represents a complete cycle of selection (Fig. 7).
- Selection response (R) depends on the total variation in the population, the bell-shaped curves in Fig. 7, the heritability (repeatability) of the trait (h^2), and the selection pressure (S) imposed. This relationship is represented in the Breeder's Equation below.

$$R = h^2 S$$

where:

$S = \mu_S - \mu$

where μ is the mean of the base population and μ_S is the mean of the selected parents

$R = \mu_0 - \mu$

where μ_0 is the mean of offspring from selected parents

Rearranging the Breeder's Equation gives heritability as:

$$h^2 = \frac{R}{S}$$

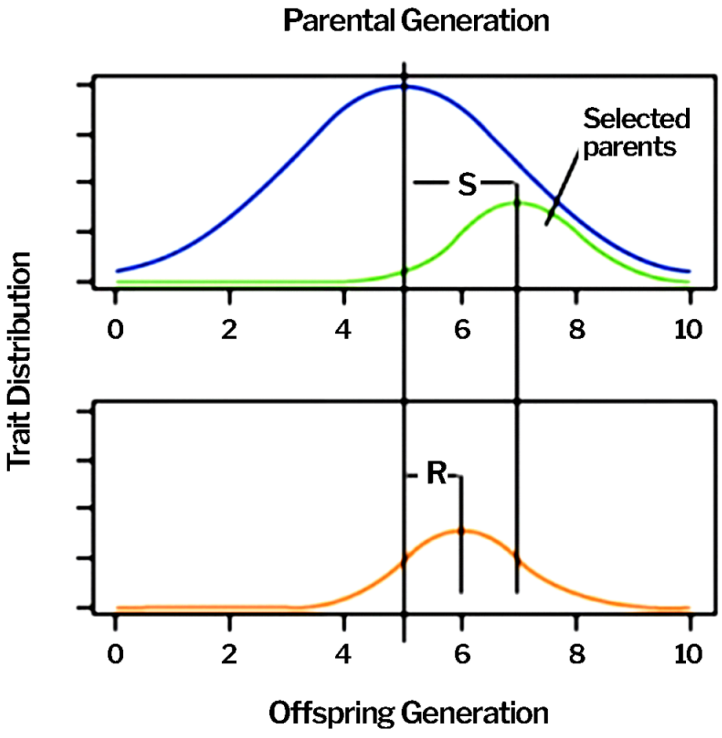


Fig. 7 Shifts in trait means from base population in parental generation and progeny generation as a result of selection.

Rate of Genetic Gain

The rate of genetic gain depends on several factors (Fig. 8): heritability of the trait, the phenotypic variability in the base population, selection intensity, and the length of the breeding cycle.

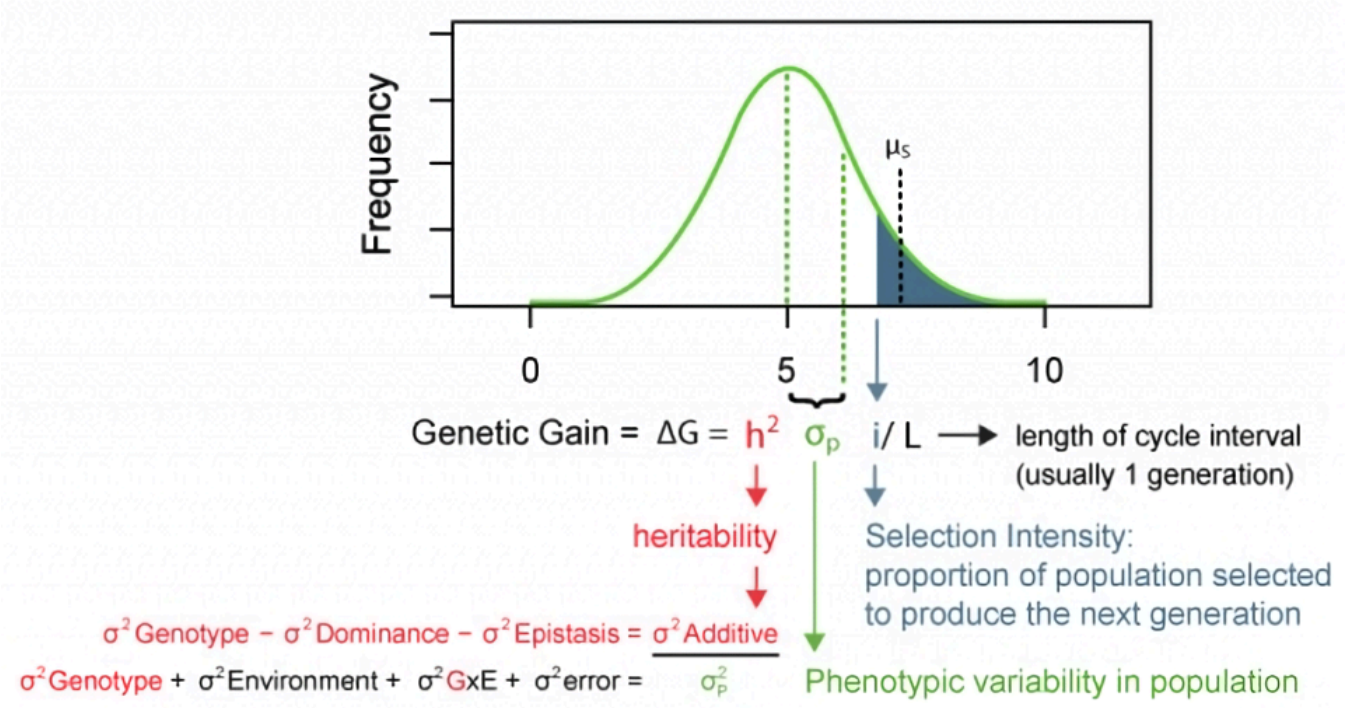


Fig. 8 Derivation of the formula for the rate of genetic gain.

Implementing the Cycle

The plant breeder must rely on a *systematic approach* to effectively increase the frequency of favorable alleles for traits of interest and maximize the rate of genetic gain so that improved cultivars are available to farmers and producers on a timely basis.

Therefore, the breeder designs a *process* by which to implement the *cycle*.

The Cultivar Development Process

The *process* of cultivar development is:

- Established by crop.
- Enabled through tools, etc.
- Focuses on practical means to develop improved cultivars.
- Typically involves several steps spanning >10 years.
- Includes all steps from crossing parents to evaluating progeny to producing volumes of seed for

distribution.

Given a finite amount of resources, the aim is to maximize genetic gain per unit of time and cost.

Core Functions in the Process

The process of developing improved cultivars involves 4 core functions that utilize different approaches to meet breeding objectives.

[Click on the **plus sign +** next to each of the core functions below to read about them.]

Note: Trait Integration is actually a special case of New Line Development and New Line Evaluation

Supporting Groups and Facilities

The process pipeline and its core functions engage supporting groups and various facilities (Fig. 9).

[Click on the **plus sign +** next to each of the core functions below to read about them.]

Multidisciplinary Engagement

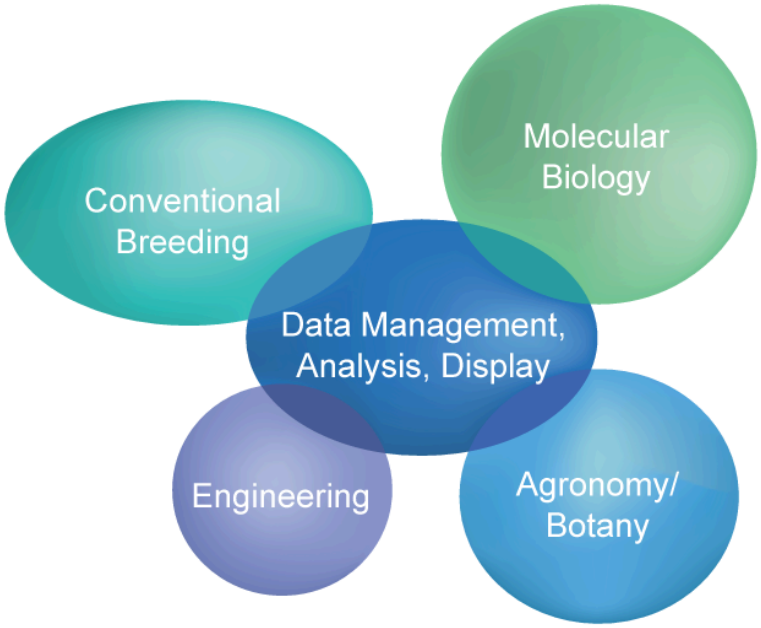


Fig. 9 The process pipeline is a multidisciplinary operation.

1. Conventional Breeding
    - Genetics: quantitative, population
    - Plant breeding methods
    - Selection theory
    - Statistics & experimental design
    - Knowledge of germplasm
    - Phenotypic evaluation
2. Molecular Biology
    - Biochemistry
    - Molecular genetic selection response (R) depends on the total variation in the population, the bell-shaped curves in Fig. 7, the heritability (repeatability) of the trait (h^2), and the selection pressure (S) imposed. This relationship is represented in the Breeder's Equation below.
        - $R = h^2 S$
          where:
          $S = \mu_S - \mu$
          where μ is the mean of the base population and μ_S is the mean of the selected parents
          $R = \mu_0 - \mu$

where μ_0 is the mean of offspring from selected parents

- Genomics
- Transformation & tissue culture
- Sequencing
- Molecular marker technologies
- Gene cloning

3. Data Management Analysis Display

   - Bioinformatics
   - Information technology
   - Information management
   - Computer programming
   - Simulation & modeling
   - Statistical and mathematical theory

4. Engineering

   - Profiling equipment
   - Analytics e.g., grain composition
   - Robotics
   - Nanotechnology

5. Agronomy/Botany

   - Plant physiology
   - Plant biology
   - Soil science
   - Pathology
   - Entomology

Timely Delivery

To succeed in delivering improved cultivars to the marketplace on a timely basis, it is *essential* to design the process to:

- Align completely with stated product targets.
- Fully integrate all aspects.

Critical Decision Points

Critical decision points involve:

- Specifying your product target and your target market region.
- Choosing optimal parents to create breeding populations.
- Developing progeny with new gene combinations.
- Evaluating progeny to identify truly superior individuals.
- Selecting progeny to advance for further evaluation and to release as improved cultivars.
- Deploying tools and technologies for greater efficiency and effectiveness.

Designing the Process: Up-front Decisions

Before beginning any activities, the breeder is faced with some important upfront decisions:

- Specify the product target.
- Define the market region for the new cultivars.
- Identify base germplasm.
- Design the breeding strategy.

Determining the Product Target

Let's talk about the product target… A breeder's specific product targets will fall out of his/her organizational mission and research team goals. Whether working for a company or for a national program, the high-level organizational mission may look something like this:

You are part of a team effort!

Organizational Mission

In a seed company, the organizational mission is translated into an overall product goal, which may be a portfolio of products for a large region. For example:

And it may include more than one crop.

Individual Breeding Program Goals

An individual breeding program will have more specific goals, centering on a particular crop.
The breeder carefully and specifically describes what is desired as a net result of the breeding process.

For example:

Note that an individual program may have more than one product target.

What is a Product Target?

The **product target** describes the "What" and the "For Where."

What *characteristics* does this product target imply?

- **Maturity group** 5 – 6 (i.e., early to medium).
- High seed yield.
- High yield stability (i.e. consistent performance across all types of environments).
- Minimal/no **lodging** (i.e., good standability at harvest).
- Minimal/no **pod shattering** ahead of harvest.
- RR1 transgenic **event** (which imparts tolerance to glyphosate herbicide).

Product Targets Indicate Specific Characteristics

Once the target characteristics are specified, the *targeted levels* of these characteristics and *the way the characteristics will be measured* must be specified. For example Table 1:

Table 1 Product Target: Characteristics, measurements, thresholds.

Characteristic	Measurement Standard	Threshold Level / Range
High seed yield	Machine harvest; seed weight at 13% moisture basis, expressed per unit of land	10% greater than Variety X
High yield stability	Use regression analysis or GGE biplot analysis	Comparable to Variety X
Lodging resistance	1-5 scale; 1=plant erect, 5=prostrate	Score ≤2
Resistance to pod shattering	Oven dry method; 10 point scale measuring percentage affected; 0=none, 1=1-10%, 10=91-100%	Score ≤1
Medium maturity	Maturity Group; day length and temperature required to initiate floral development; full range includes Group 000 to Group 9	MG 5-7
RR1 event (Roundup Ready 1)	Integrate 40-3-2	Pre-determined level of glyphosate tolerance

The target levels for the target characteristics become the thresholds that will be used in selection.

For value-added traits like RR1, achieving the desired level of trait expression is typically a function of integrating the particular transgenic event through either backcross or forward breeding.

Product Targets Indicate a Specific Market Region

What market region is specified in this Product Target?

In this case, the market region is “dryland and irrigated areas of South Africa involving corn production.” Because the purpose of soybean in corn rotation is to fix nitrogen in the soil, certain bacteria must be present in the soil to facilitate this activity. Soybean works together with Rhizobia and other bacterial species to convert atmospheric nitrogen to a form readily usable by plants, presumably the corn crop in the following year. To ensure that nitrogen-fixing strains of bacteria are present, farmers may inoculate the soil. In this case, use of the inoculant “Rhizobia strain WB74” is specified as a production management practice.

These details speak to the *population of environments*, that is, the target market in terms of geography, production management system, season, maturity zone, altitude, etc.:

- Locations used as testing environments must be representative in terms of locations, planting dates, farmer practices, soil types, etc.
- Think of your testing environments as "samples" from the population of environments.

Measurements and Thresholds

Let's look further…

The product target for an improved soybean variety for South Africa requires certain characteristics:

- Maturity group 5 – 6 (i.e., early to medium).
- High seed yield.
- High yield stability (i.e., consistent performance across all types of environments).
- Minimal/no lodging (i.e., good standability at harvest).
- Minimal/no pod-shattering ahead of harvest.
- RR1 transgenic event (the "Roundup Ready 1" event imparts tolerance to glyphosate herbicide).

And, with these requirements outlined, the breeder can designate the way the characteristics will be measured i.e., what "screens" will be used to evaluate performance for the trait. For example, to measure resistance to pod shattering, a protocol is needed for the "oven dry method" of evaluation.

The breeder can set the threshold levels for identifying and selecting superior progeny (Table 1).

Identifying Base Germplasm

The breeder also considers what germplasm to use as parents in creating progeny with useful, new gene combinations. Sources representing a high frequency of favorable alleles for the traits to be improved are needed.

Designing the Breeding Strategy

The breeder can even begin to consider breeding strategies to be used to assemble the suite of characteristics in one line.

Breeding strategies incorporate testing schemes and selection criteria, and also consider the following:

- mode of reproduction.
- selection intensity (i.e. how many individuals will be selected at each evaluation step) and order in which selection for particular traits is implemented.
- breeding methods.
- systems and facilities available.
- experimental designs for performance evaluation.
- predicted response to selection.

After all, the total package of characteristics is what will constitute the new, improved variety.

Types of Cultivars

In addition to the required characteristics and specific market, the product target indicates the type of cultivar to be developed.

There are several types of cultivars:

- A pure line variety is **homozygous** and **homogeneous**. It can be considered an inbred line. Example: soybean, pea.
- A hybrid is the result of crossing two genetically different lines. It is designed to exploit "heterosis", or hybrid vigor, which may be expressed as increased yield and more robust plant health. Example: maize.
- An open-pollinated variety (OPV) is similar to a random mating population wherein cross-pollination occurs due to wind, insects, birds, and other natural mechanisms. Example: carrots.
- A synthetic is a population of cross-pollinated plants, typical of crop species that are self-incompatible for self-pollination. Example: alfalfa.
- A blend is a mixture of genotypes intended for genetic diversity to promote yield stability. The blend may comprise different types of disease resistance, slightly different maturities, or varying levels of winter hardiness. Example: wheat.
- Clonally propagated cultivars are genetically identical to a "mother plant" that is the result of cross-pollination. Often these species are polyploids. Example: potato.

Factors to Consider With Type of Cultivar

The reproductive system of the crop in question, its life cycle, and its **ploidy level** may dictate the types of cultivars that can be developed. For example:

- Obligate out-crossing species cannot produce pure line cultivars.
- Perennial crops include fruit tree species, many of which are produced from highly heterozygous parents; these are often reproduced and distributed as clones.
- Some seedless cultivars are the result of hybridizing parents of different ploidy numbers. Case and point: seedless watermelon results from crossing tetraploid with diploid watermelon to produce a triploid cultivar, which is used in production fields with a diploid pollinizer to produce sterile (seedless) fruit.

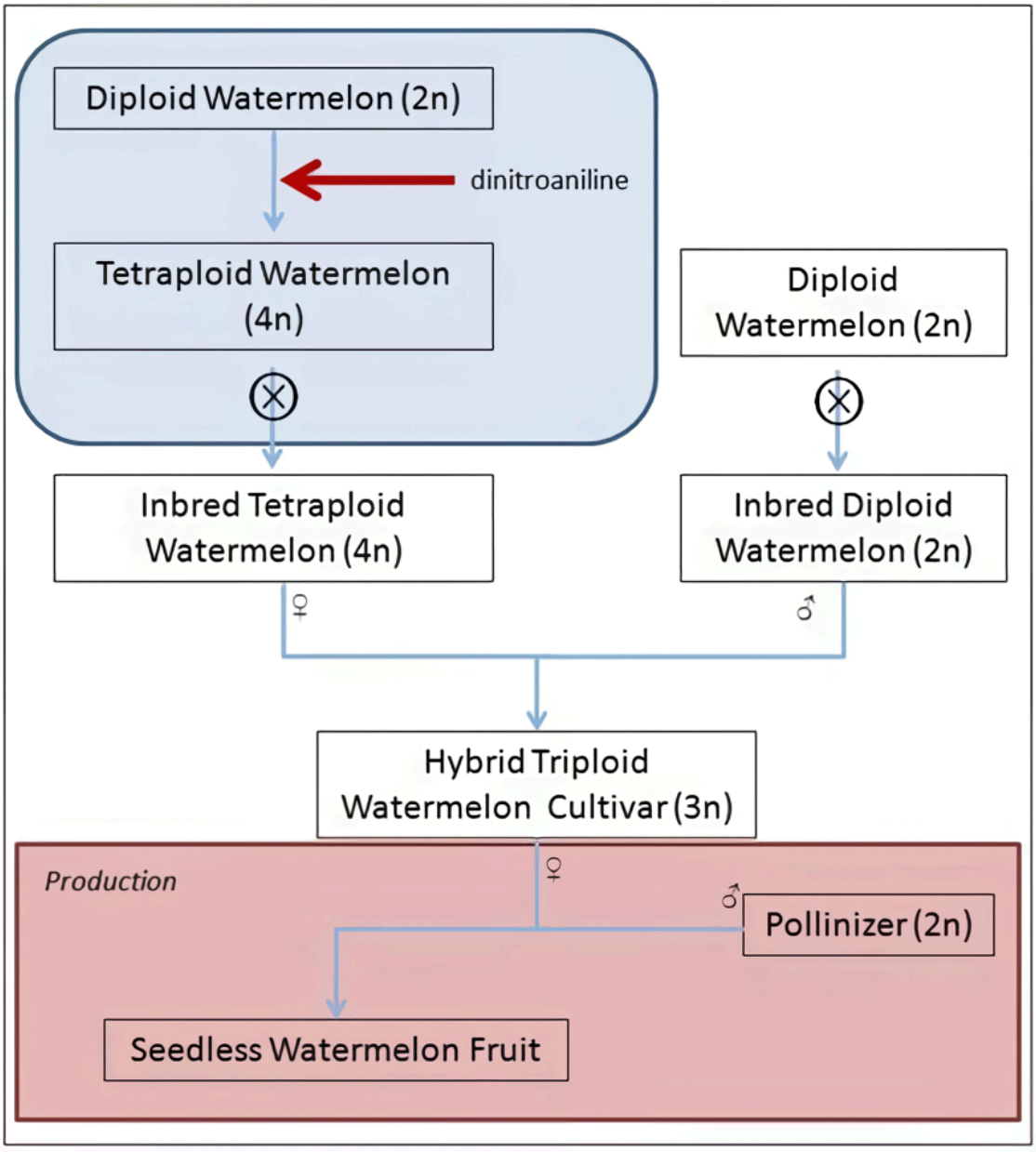


Fig. 10 Production of seedless triploid watermelon. Image Credit: Geoffrey Meru, Department of Horticulture, University of Georgia.

Hybrid Cultivars

Crops displaying heterosis may lend themselves to hybridization as a means to spike yields. Some crops that have transitioned from pure lines or OPVs to hybrids, or are in process, include maize, rice, sorghum, pearl millet, wheat, barley, sunflowers, cucumbers, tomatoes, melons, squash, and others. One essential requirement for hybrid cultivars is a dependable, cost-efficient means to control pollination in producing the hybrid seed (e.g. cytoplasmic male sterility, mechanical means to remove male flowers, environmentally- or chemically-induced male sterility, or a **dioecious** crop species).

In addition, farmers must be willing to purchase fresh seed each year (saved seed will not display the same levels of hybrid vigor).

Fig. 11 An early maturing grain sorghum hybrid exhibits tolerance through droughty conditions. Photo by Susan O'Shaughnessy, courtesy of USDA.

Process Efficiency

To recap...

The *process* of cultivar development becomes the mode and mechanism to implement the *cycle* of cultivar improvement effectively.

With the product target defined and the target market specified, *the process must be designed to maximize the rate of genetic gain* in order to accelerate the release of improved cultivars (Fig. 4).

$$\Delta G = \frac{ih^2\sigma_p}{L}$$

Example: Commercial Soybean Improvement Program

Let's look at an example of a commercial soybean breeding program. Note that this program serves a market region with one growing season (summer):

Table 2 New Line Development and New Line Evaluation in a commercial breeding program for soybean. Adapted from Bernardo, 2010.

SEASON*	ACTIVITY
Winter 0	Make breeding crosses
Summer 1	Self or BC each population
Winter 1a**	CONSISTENT use of SUBSCRIPTS Grow 200 F_2 or BC_1 populations (i.e. S_0 generation) that have been formed in previous years Advance the S_0 plants to the S_1 generation by a modified single-seed-descent method, retaining single pod (instead of a single seed) with 2-3 seeds; bulk by population
Winter 1b	Plant S_1 seed bulk; Select 200-500 plants (~350 per population) and save selfed (i.e. S_2) seed
Summer 2	Yield trials for 70,000 S_2 families (across all populations) in unreplicated trials at 1-2 locations; Select the best 5,000 based on yield performance; Save S_3 seed from the trials
Summer 3	Yield trials for 5,000 S3 families at 3-5 locations; Select the best 200 based on yield performance; Save S_4 seed from the trials
Summer 4	Yield trials for 200 S_4 families at 15-25 locations; Select the best based on yield performance; code selected as 'experimental' lines; Save S_5 seed from the trials
Winter 4	Increase seed of experimental lines
Summer 5	Yield trials of experimental lines at 20-40 locations On-farm strip tests (i.e.150-300 m^2 plots) at 20-100 locations
Summer 6	Yield trials of 'advanced' lines at 20-50 locations On-farm strip tests (i.e. 150-300 m^2 plots) at 30-500 locations
Fall 6	Release 0-5 new varieties

*Summer represents the main growing season, winter denotes off-season activities; the number after season indicates the year in the development pipeline

**Winter nurseries may be grown back-to-back in the same winter season

Notable Components

Components of the example soybean program:

- ≥7 years from the first cross to the commercial release of a new pure line variety.
- ~200 populations initiated per year.
- 500-1000 families per population.
- The number of new lines decreases as number of locations increases.
- 5+ years of yield testing before launch.

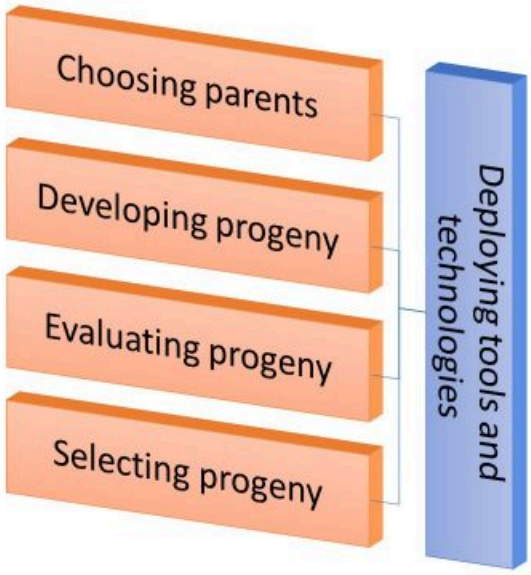
Course Roadmap

The remainder of this course will focus on how to design and implement the "process" of cultivar development so as to increase the likelihood of meeting your specified product target and identifying truly superior progeny among your breeding populations.

We will examine key decision points in the process pipeline and weigh the issues related to choices in:

- Choosing parents (to create breeding populations).
- Developing progeny (with new gene combinations).
- Evaluating progeny (to identify truly superior individuals).

- Selecting progeny (to advance for further evaluation and to release as improved cultivars).
- Deploying tools and technologies (for more efficiency and effectiveness) at various steps in the process pipeline.



Course Organization

In the remainder of this course,

- We will first walk through the "process" related to development of an improved pure line variety (the simplest example).
- We will then look at how the process might change if a hybrid cultivar is the target or if a clonally-propagated cultivar is the target.
- We will take a closer look at Trait Integration, which is a special case of New Line Development/New Line Evaluation where one parent serves only as the source of an important trait (e.g. preparing and utilizing sources of a value-added trait).
- We will explore ways to optimize the process pipeline to maximize gain from selection.
- Finally, we will examine aspects of product launch and producing volumes of seed for release and distribution of a new, improved cultivar.

References

Bernardo, R. 2010. Breeding for Quantitative Traits in Plants. 2nd edition. Stemma Press, Woodbury, MN.

Mumm, R.H. 2013. "A look at seed product development with genetically modified crops: Examples from maize." J. Agricultural and Food Chemistry 61(35): 8254-8259. DOI: 10.1021/ jf400685y.

Ray, D.K., N.D. Mueller, P.C. West, J.A Foley. 2013. "Yield Trends Are Insufficient to Double Global Crop Production by 2050." PLOS ONE 8(6): e66428. doi:10.1371/journal.pone.0066428

How to cite this chapter: Mumm, R.H. (2023). Chapter 1. Plant Breeders Rise to the Challenge of Feeding the World. In W. P. Suza, & K. R. Lamkey (Eds.), *Cultivar Development*. Iowa State University Digital Press.

Chapter 2: The Process of Cultivar Development: Pure Line Variety

Rita H. Mumm

The process of cultivar development for pure line varieties involves choosing parents, creating progeny and materials for testing, and evaluating and selecting desirable lines. The details of each step are described below.

Choosing parents

Cycle to Process

The *process* of cultivar development becomes the mode and mechanism to effectively implement the *cycle* of cultivar improvement (Fig. 1).

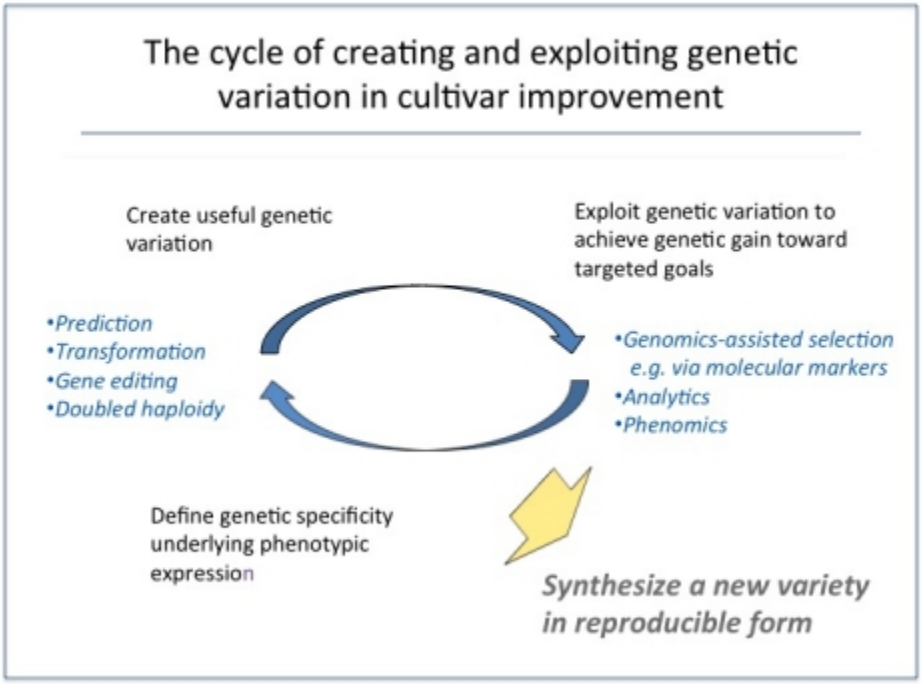


Fig. 1 Inclusion of modern tools in the cycle of creating and exploiting genetic variation in cultivar improvement. Adapted from Mumm, 2013.

Process to Pipeline

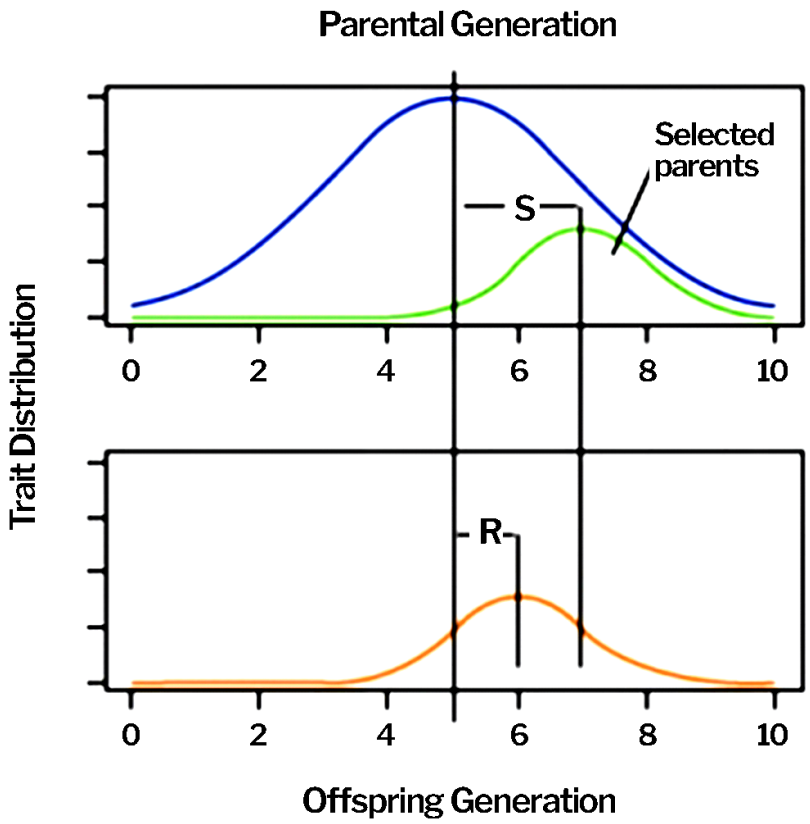


Fig. 2 Shifts in trait means from base population in parental generation and progeny generation as a result of selection.

The process is fashioned to increase the frequency of favorable alleles for the traits specified in the product target… and this goal is reflected in the way that parents are chosen, outstanding progeny are identified, and new cultivars are created (Fig. 2). *R* is selection response, *S* is selection pressure, and h^2 is heritability in the Breeder's Equation

$$R = h^2 S$$

and

$$h^2 = \frac{R}{S}$$

The process is designed to produce a pipeline of improved cultivars, maximizing response to selection (*R*) and the rate of genetic gain (ΔG).

Maximizing Gain

Looking more closely at the equations, *R* can be written a number of ways:

$$\begin{aligned} R &= ih^2\sigma_p \\ &= i\frac{\sigma^2}{\sigma_p^2}\sigma_p \\ &= i\frac{\sigma_a}{\sigma_p^2}\sigma_p \\ &= ih\sigma_p \end{aligned}$$

where:
h^2 is narrow sense heritability,
σ_P is phenotypic standard deviation,
σ_P^2 is phenotypic variance,
i is selection intensity,
σ_a is additive genetic standard deviation,
σ_a^2 is additive genetic variance, and
h is accuracy of selection (the square root of h^2).

And can be expressed as:

$$\Delta G = \frac{ih^2\sigma_p}{L}$$

where *L* is the length of a breeding cycle.

Considerations in Choosing Parents

In choosing parents, consider:

- best performers for traits of interest (i.e. contributors of favorable alleles).
- genetically diverse contributions of favorable alleles (i.e. genetic diversity so as to boost the overall frequency of favorable alleles and the number of loci with the favorable allele present).

Germplasm Banks

Base germplasm can be accessed through the many germplasm banks around the globe. For soybean, there are 189 such repositories, the major germplasm banks listed by the FAO (2010) are listed in Table 1 (extract from Table 2 in Jacob et al, 2016).

Table 1 Listing of Major Germplasm Repositories for Soybean.

Germplasm Bank	No. Accessions	No. Advanced Cultivar Lines
ICGR-CAAS (Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences)	32,021	n/a
SOY (Soybean Germplasm Collection, United States Department of Agriculture, Agricultural Research Services)	21,075	84
RDAGB-GRD (Rural Development Administration National Institute of Agricultural Biotechnology-Genetic Resources Division, Republic of Korea)	17,644	176

Primary, Secondary, and Tertiary Gene Pools

Germplasm banks may contain materials from the primary, secondary, or tertiary gene pools denoted as GP-1, GP-2, and GP-3, respectively (as defined by Harlan and de Wet, 1971) (Fig. 3).

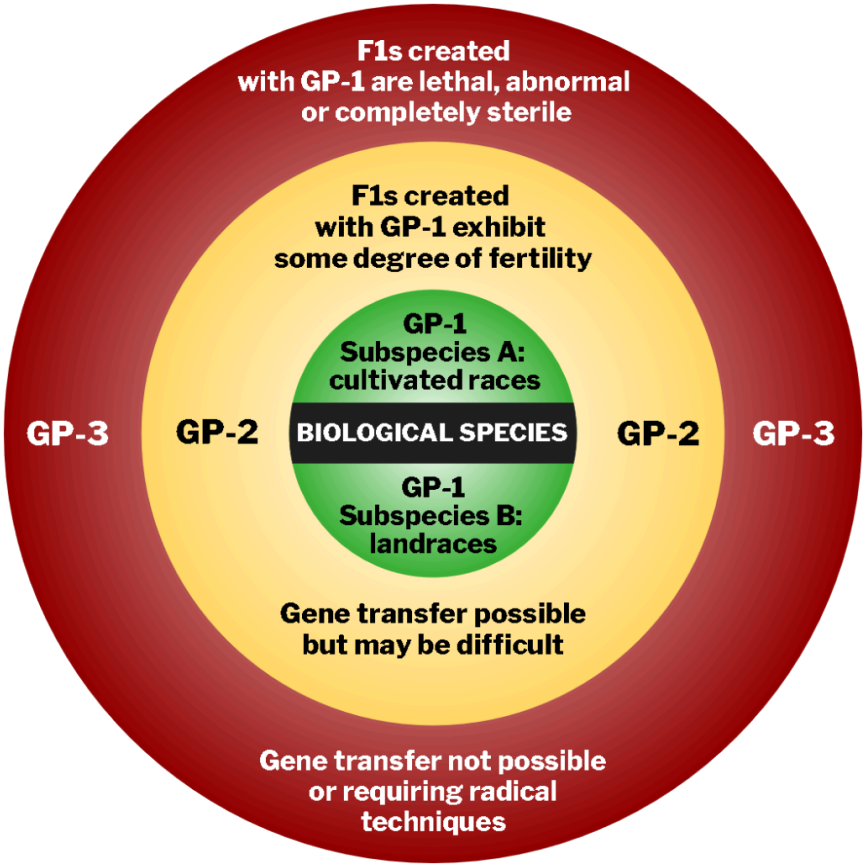


Fig. 3 Primary, secondary, and tertiary gene pools.

Sources of Parental Germplasm

Potential sources of parental germplasm include:

- Current cultivars.
- Elite breeding lines.
- Acceptable breeding lines with superiority in one or a few characteristics (e.g., genetic stocks).
- Plant introductions, landraces.
- Related species such as wild relatives.

Which of these source types could be utilized to create a “good x good” cross?

What is a “good x good” cross?

A “good x good” cross is ideal!

Both parents possess favorable alleles for each character, giving a high likelihood that at least some of the progeny will exceed the performance of either parent for each character.

What factors determine that likelihood?

- The expected frequency of desired genotypes in the breeding population for each character.
- The number of traits being selected.
- The inheritance of the character to be improved.
- The number of progeny evaluated.

Favorable alleles contributed by each parent are distinctive (i.e., parents contribute different favorable alleles).

Ideal Parents Possess Favorable Alleles

To determine if prospective parents possess favorable alleles for traits of interest, consider:

- High mean phenotypic performance
- In a self-pollinated crop like soybean, the **mid-parent value** is the best indicator of performance of progeny from a prospective cross.
- Favorable molecular marker profile (e.g., mapped QTLs)
- Estimated breeding value

Ideal Parental Combinations are Genetically Diverse

To determine if prospective parents are genetically diverse, utilize:

- Pedigree analysis.
- Geographic inference.
- Cluster analysis based on molecular marker profile.

Looking for evidence that favorable alleles for traits of interest are not the same alleles between prospective pairs of lines to use in making a breeding cross

Assessing Genetic Diversity Based on Molecular Marker Profile

Genetic diversity of prospective parents can be assessed using cluster analysis based on molecular marker profiles.

Lines are genotyped for a set of markers and marker alleles are scored. Each pair of lines is compared for each marker locus to compute an overall estimate of similarity, which is later converted to an estimate of dissimilarity.

Cluster analysis provides a visual depiction of the relative dissimilarity among the lines, which is referred to as a dendrogram (Fig. 4).

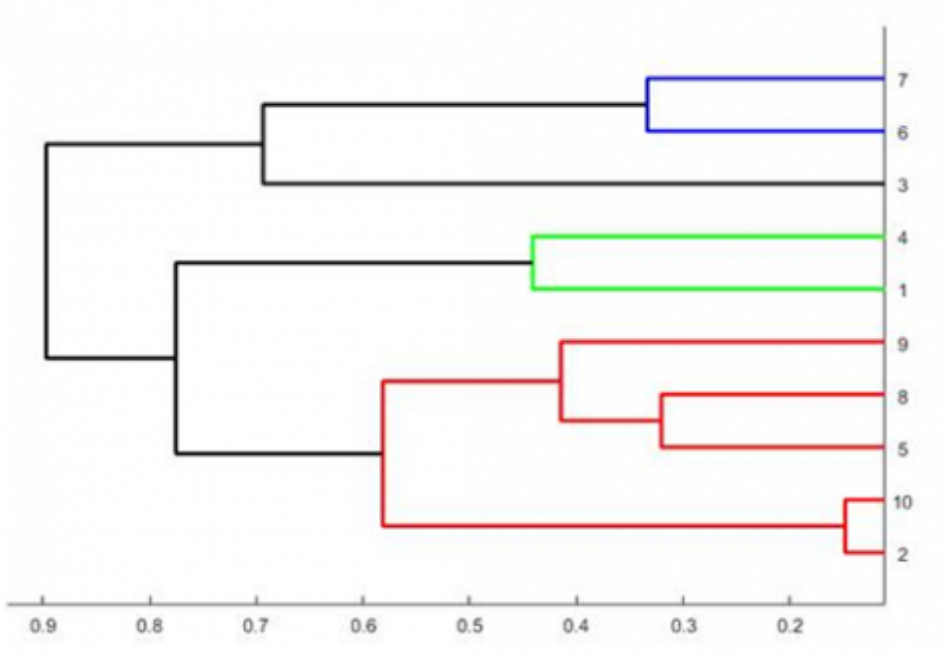


Fig. 4 Dendrogram illustration showing clusters of various levels of diversity.

Tips for Cluster Analysis

In performing cluster analysis based on marker data:

- Consider the type of marker to sample the genome and the number of markers needed for adequate coverage.
- Choose a method for estimating dissimilarity.
- Choose a method for joining clusters.
- Include some familiar lines of known background in the analysis to serve as “anchors”.
- Remember that the output is relative to the lines included in the analysis, not absolute.

Clustering Based on Pedigree Records

Cluster analysis can also be used to represent genetic diversity graphically, based on complete and accurate pedigree records. The genetic diversity of 38 soybean lines prominently used as parents to develop the current U.S. commercial soybean germplasm base is depicted in the dendrogram below. The results are from cluster analysis based on dissimilarity estimates computed from the coefficient of parentage (Fig. 5).

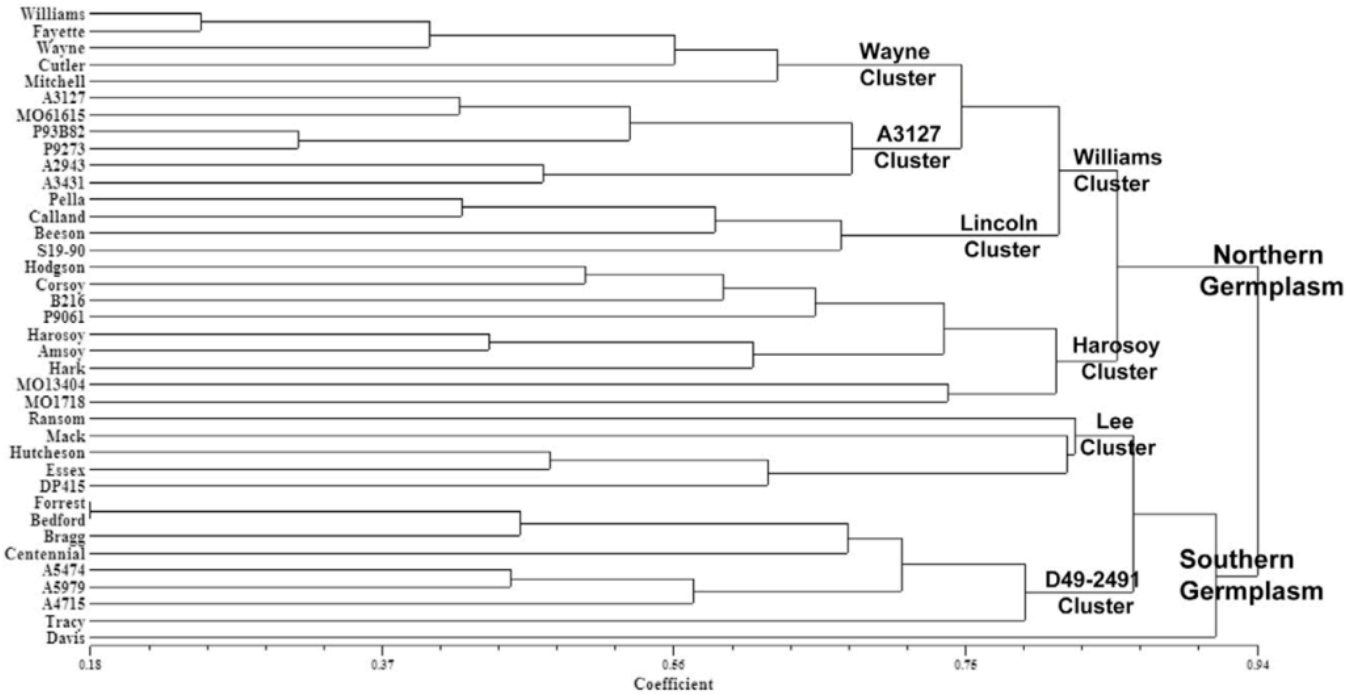


Fig. 5 Dendrogram of 38 soybean lines from Mikel et al. 2010.

Suitable Parents for a Stated Product Target

Consider the following product target…

The product target can be dissected to specify the desired characters, their measurement standards, and threshold levels. For example, Table 2 below.

Table 2 Soybean product target: Characteristics, measurements, and thresholds.

Characteristic	**Measurement Standard**	**Threshold Level/ Range**
High seed yield	Machine harvest; seed weight at 13% moisture basis; expressed per unit of land	10% greater than Variety X
High yield stability	Use regression analysis or GGE biplot analysis	Comparable to Variety X
Lodging resistance	1-5 scale: 1=plant erect, 5=prostrate	Score ≤ 2
Resistance to **pod shattering**	Oven dry method; 10 point scale measuring percentage affected: 0=none, 1=1-10%, 10=91-100%	Score ≤ 1
Medium maturity	Maturity Group; day length & temperature requirement to initiate floral development; full range includes Group 000 to Group 9	MG 5-6
RR1 Event (Roundup Ready 1)	Integrate 40-3-2	Pre-determined level of glyphosate tolerance

Describe the characteristics of parents comprising a “good x good” cross to begin New Line Development in terms of:

- Types of germplasm.
- Traits exhibited in each prospective parent and levels of each trait.

What if the ideal parental combination is not feasible?

Assessing Prospective Parents for Favorable Alleles

Whereas cluster analysis highlights genetic diversity, it is not an indicator of favorable alleles for the traits of interest in the prospective parents.

Mean performance provides insights as to whether favorable alleles are present… what we really want to know is the *breeding value* of prospective parents.

Parents pass on their genes, not their genotype or phenotype, to their progeny.

Breeding Value Defined

Breeding value (BV) is…

- The value of an individual, judged by the mean value of its progeny.
- The value of genes, not genotype, passed on to progeny.
- The sum of the average effects of genes in an individual.

Falconer (1989) defines BV as the mean deviation from the population mean of individuals who received that allele from one parent, the other allele coming at random from the population.

On a single locus example, the sum of the average effect of alleles:

If A_1 = 10 and A_2 = –10
then BV of A_1A_1 = 20
BV of A_1A_2 = 0
BV of A_2A_2 = –20

Estimated Breeding Value

Estimated breeding value (EBV) is a function of narrow sense heritability (i.e., only additive genetic variance is included).

$$EBV = h^2 P$$

Phenotype (P) is expressed as deviation from the population mean (μ).

Example:

If P = +3.0 units (i.e. standard deviations) and h^2 = 0.33,

Then EBV = +3.0 x 0.33 = +1.0 unit

EBV Predicts Progeny Performance

Estimated breeding value is…

- A statistical prediction of the relative genetic merit of individuals as parents.

Crow (1986) explains that the predicted phenotype of the progeny is the average of the breeding values of the parents.

- Used to rank available candidates for use in developing new breeding populations.
- Computed through methods such as:
  – BLUP (Best Linear Unbiased Prediction).
  – GS (Genomic Selection).

Best Linear Unbiased Prediction

Best Linear Unbiased Prediction = **BLUP**

- **B**est: Minimum error.
- **L**inear: Estimates are linear functions of the data.
- **U**nbiased: The average value of the estimate is equal to the average value of the quantity being estimated.
- **P**rediction: Estimates of random effects are typically referred to as 'predictors' (estimates of fixed effects are called 'estimators').

Origin of BLUP

BLUP originated in animal breeding, where choice of parents is critical to maximizing the use of an expensive sire. The progeny resulting from the breeding cross live for years and comprise the herd.

In plant breeding, we choose many breeding pairs and we can discard any breeding populations we don't like!

BLUP Model

Phenotype = Environmental Effects + Genetic Effects + Residual Effects

$$y_{ij} = \mu_i + g_i + e_{ij}$$

where:
y_{ij} is the jth record observed for the ith line
μ_i is the identifiable nonrandom (fixed) environmental effects such as year, location
g_i is the sum of the additive, dominance, and epistatic effects of the ith line
e_{ij} is the sum of the random environmental effects from the jth record of the ith line

Partitioning Additive Effects

Partitioning the additive effects from the total genotypic effects, the model can be written as:

$$y_{ij} = \mu_i + (g_a)_i + (g_d)_i + (g_e)_i + e_{ij}$$

$$y_{ij} = \mu_i + (g_a)_i + e^*_{ij}$$

with e^*_{ij} including all nonadditive genetic effects as well as error.

Dealing With Fixed Effects

The concept of "fixed effects" stems from the principle of *fair* comparisons. Although phenotypes and breeding values are often expressed as deviations from group means, it is desirable to make comparisons on an equitable basis.

For example, it would not be fair to compare yields in a wet year to yields in a dry year.

To compare individuals on a comparable basis, phenotypes are adjusted for known fixed effects.

The Mixed Model Equation

The mixed model equation partitions fixed and random effects:

$$Y = X\beta + Z\mu + e$$

where:
Y = the vector of observed values,
β = the vector of fixed effects,
μ= the vector of random genetic effects (the EBVs),
e = the vector of residuals (error),
X and Z are design matrices that relate elements in β and μ to elements in Y.
Thus, both breeding values and fixed effects can be estimated.

Solving the Mixed Model Equation

To estimate BVs, solve for:

$$\begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\alpha \end{pmatrix} \begin{pmatrix} \beta \\ \mu \end{pmatrix} = \begin{pmatrix} X'Y \\ Z'Y \end{pmatrix}$$

where:
A^{-1} is the inverse of the matrix of genetic relationships among individuals and,
$\alpha = \dfrac{(1 - h^2)}{h^2}$ where h^2 is the narrow-sense heritability for the trait under analysis.

Note: Molecular markers can be used to estimate genetic relationships among lines (instead of pedigree information).

Advantages of BLUP over Mid-Parent Value

In work by Panter and Allen (1995), BLUP was more effective than mid-parent value (MPV) for identifying superior parental combinations and demonstrated higher probability of producing superior progeny.

One clear advantage of BLUP over MPV is that the potential of a particular pair of lines as parents could be predicted with BLUP even when no performance data on the lines was available, only performance data on relatives of the lines was available.

For more information, see: Panter, D.M., and F.L. Allen. 1995. Using Best Linear Unbiased Predictions to Enhance Breeding for Yield in Soybean: I. Choosing Parents. Crop Science 35: 397-405.

Advice for Use of EBVs

Tips for use of EBVs:

- EBVs are best used as one piece of information along with all the rest of the knowledge base about prospective parents.
- EBVs are strictly a function of the data used to generate the estimates. You may wish to add new phenotypic data as available (more locations, more years, more related lines) and rerun BLUP.

To distinguish BV from genotypic value:

Breeding value is the expected phenotype of an individual's progeny and includes only additive effects, whereas
Genotypic value is the expected phenotype of an individual given its genotype; includes additive and non additive effects.

What is Genomic Selection?

Genomic Selection is comprised of methods that use genotypic data across the entire genome to predict performance of any quantitative trait with high accuracy (Fig. 6).

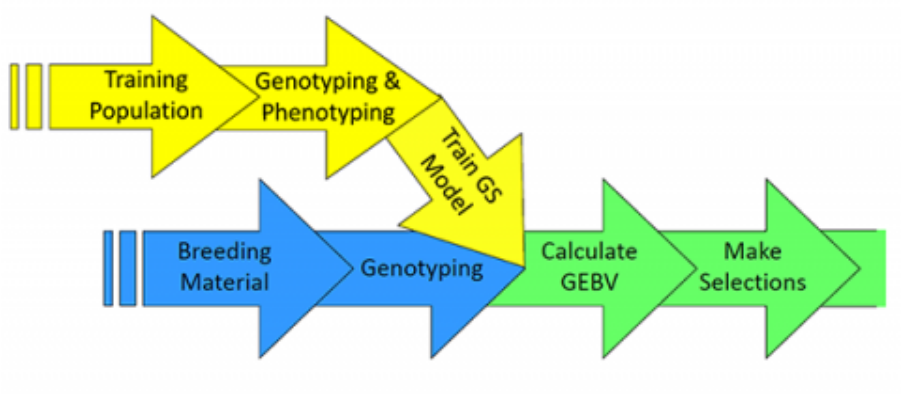


Fig. 6 Genomic selection process; from Heffner et al. 2009.

A collection of representative individuals (i.e., Training Population) is phenotyped and genotyped to "train" a model with estimated effects for a genome-dense set of molecular markets for a particular quantitative trait.

Genomic Estimated Breeding Value

How is Genomic Selection applied to choose among prospective parents?

- Estimates of effects are computed for each marker allele using an appropriate training population.
- **Genomic Estimated Breeding Values** (GEBVs) are calculated for prospective parent lines based on the model. The GEBV is essentially a sum of the effects across the individual's genome.
- Individuals with the highest GEBVs are used as parents to create the next generation. They are also advanced to varietal testing for possible release as new cultivars, bypassing preliminary testing (Fig. 7).

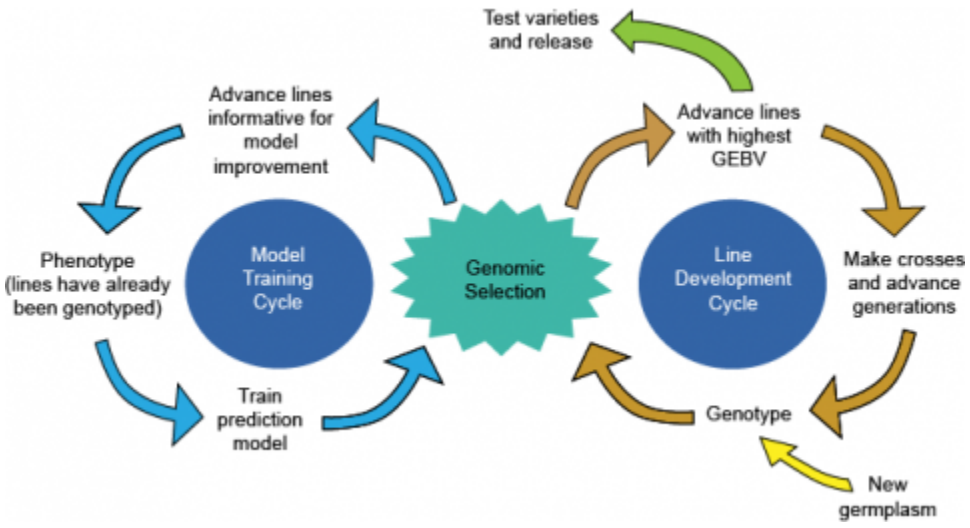


Fig. 7 Diagram of a genomic selection breeding program. Adapted from Heffner et al. 2009.

Use of Suboptimal Parents

So far, we have considered the scenario when two elite lines are suited as parents to supply favorable, diverse alleles for the traits specified in a product target. This is the ideal.

Now, let's examine the situation where you are not able to work with a "good x good" cross.

What are some of the reasons why suboptimal crosses may be the only option available to accomplish a particular product target?

Moderating the Effect of a Non-elite Parent

With a "good x not-so-good" cross, the breeder can modulate the effect of the non-elite parent through backcrossing to the elite parent before creating progeny for evaluation.

This can take different forms:

1-3 backcrosses to the elite parent to access potentially favorable genes for a quantitative trait from a lesser elite line, a **genetic stock**, an elite but unadapted cultivar, or a wild relative (Fig. 8).

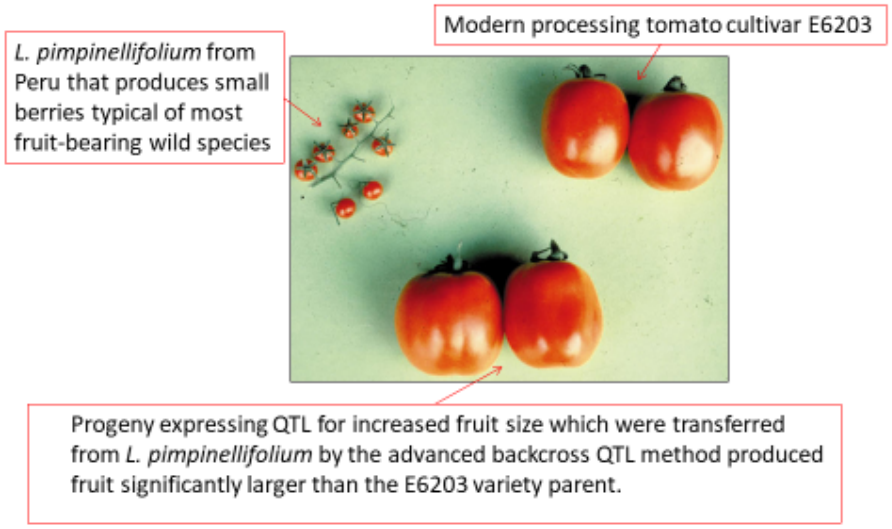


Fig. 8 Example of yield improvement in tomato from Tanksley et al. 1996

≥6 backcrosses to introgress a single gene into an elite cultivar to create an **isoline** that performs equivalently in every way as the elite parent except for the integrated new trait (e.g., trait integration with disease resistance) (Fig. 9).

Fig. 9 The lettuce mosaic virus, which can cause a mottling pattern on the leaves, is controlled by a single recessive allele (mo1) (Irwin et al. 1999). Photo by Dennis H. Hall.

The Role of Backcrossing

Backcrossing (ahead of self-pollination) allows the breeder to determine the desired "dosage" of the non-elite parent genome.

With each repeated crossing to the **recurrent parent** (i.e. elite parent), the amount of germplasm from the **non-recurrent parent** (i.e. non-elite parent) is reduced by half (Fig. 10).

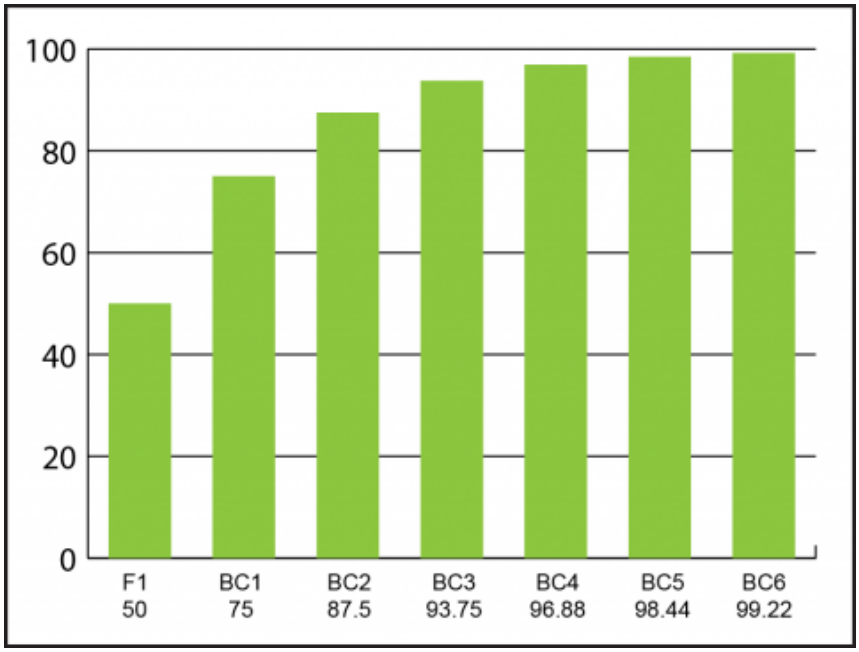


Fig. 10 Expected percentage of RP germplasm recovered with each cross to the RP.

$$\%RP(n) = 1 - \left(\frac{1}{2}\right)^{n+1}$$

where:
$\%RP$ is the percentage of recurrent parent germplasm recovered, and
n is the backcross generation number.

Choice of Parents Affects Choice of Breeding Methods

If a suboptimal cross is the only available option to meet a specific product target, the approach to breeding strategies will look different than the ideal situation where two elite parents can provide the favorable alleles needed.

Bottom line: Choice of parents affects the choice of breeding methods!

Types of Parental Crosses

Although 2-parent combinations are the most widely used, breeding populations can be developed using multiple parents. For example:

- A 3-parent breeding population is formed by mating a 2-parent population to a third parent: (P1 x P2) x P3. The plants of the 2-parent population may be at F_1 (as indicated by this pedigree)

or some later stage of inbreeding.

- A 4-parent population, often referred to as a "4-way cross", can be formed in two ways:
  - (P1 x P2) x (P3 x P4) where each parent contributes ~25% of the alleles
  - [(P1 x P2) x P3] x P4 where P1 and P2 contribute ~12.5% germplasm, P3 contributes ~25%, and P4 contributes ~50%.
- A complex breeding population is formed using >4 parents.

In each case, the "dosage" of each parent is determined by the way the population is formed.

Complex Parental Crosses

A complex breeding population can involve five to hundreds of parents! This comes with advantages and disadvantages (Table 3).

Table 3 Advantages and disadvantages of complex parental crosses.

Advantages	Disadvantages
The greater the number of parents: • The greater the number of possible alleles in the population for each locus • The greater the probability that ≥ 1 parent has the most favorable allele at each locus • The greater the probability that homozygous parents will have different alleles at ≥ 2 linked loci.	Expected mean peformance of the progeny is still the mean of the parents for self-pollinated crops. If some parents are not elite for traits of interest, performance is not likely to exceed the performance of the best parent for each trait.
Greater probability of heterozygosity at multiple loci.	The greater the number of parents, the more generations required for inter-mating, which can lead to significant time expansion of a breeding cycle compared to 2-parent populations.
More allelic combinations possible!	Might not be a efficient as multiple 2-parent populations.

Creating Progeny and Materials for Testing

Mating Design Determines the Selection Unit

After deciding on the parental combinations, the breeder must make decisions about the progeny to be produced and evaluated (Fig. 1). The **mating design** employed determines the "selection unit."

Considerations in Choosing a Mating Design

Factors to consider in choosing a mating design:

- Mode of reproduction (mating flexibility).

- Specific objectives
  - Understanding gene action
  - Developing estimates of genetic variance components and heritability which are necessary to designing the breeding process structure
  - Choosing among breeding methods
  - Predicting gain from selection
  - Making selections and developing an improved cultivar
- Reliability of the estimates.

Rule: use the simplest design which meets your needs.

For self-pollinated crops, full sib designs are common.

Mating Design Details

Specifics to be determined in choosing a mating design:

- Types of progeny (e.g., full sibs, half sibs, doubled haploids).
- Number of types of progeny and scheme for how parents will be organized.
- Total number of progeny.

Mating Design to Estimate Genetic Variance

Estimation of generic variance requires four steps:

1. One or more types of progeny developed.
2. Progeny evaluated in set of environments.
3. Variance components estimated from mean squares in the ANOVA.
4. Variance components interpreted in terms of covariances between relatives.

The Case of the 1-Factor Design

With the use of a single type of progeny, the ANOVA includes one component of variance for progeny and one covariance between relatives. The Expected Means Square table provides a way to estimate the component of variance. This estimate can then be interpreted in terms of the covariance of relatives it represents.

For example (Table 4).

Table 4 Expected Means Square table.

Source of Variation	d.f.	EMS
Reps	$k-1$	n/a
Progenies	$p-1$	$\sigma^2 + k\sigma^2_{prog}$ [Design component of variance]
Error	$(k-1)(p-1)$	σ^2
Total	$n-1$	n/a

If progeny are clonally reproduced, then:

$$\sigma^2_{prog} = \text{Cov Identical Twins} = \sigma^2_G$$

If half sib families, then:

$$\sigma^2_{prog} = \text{Cov HS} = \frac{1}{4}\sigma^2_A$$

If full sib families, then:

$$\sigma^2_{prog} = \text{Cov FS} = \frac{1}{2}\sigma^2_A + \frac{1}{4}\sigma^2_D$$

Other Mating Designs

Mating designs are classified according to the number of factors, parents, and modalities or a combination of these factors (Cockerham, 1963).

Common mating designs of interest:

- 1-Factor design.
- 2-Factor designs can estimate dominance and additive variances:
  - Design I: hierarchical, nested
  - Design II: factorial design—cross-classified, parents must be inbreds
  - Diallel: used to estimate GCA and SCA
- 3-Factor Designs have at least one grandparent in common and can estimate additional variance components (e.g., epistatic)
  - Design III: can screen out psuedo-overdominance.

Level of Inbreeding

Another decision involves the level of inbreeding… will selection be initiated with segregating materials or fully homozygous lines?

Evidence to support Early Generation Testing (Fig. 11) is the strong correlation between early generation performance and late generation performance for yield and other quantitative traits, even with lowly heritable traits.

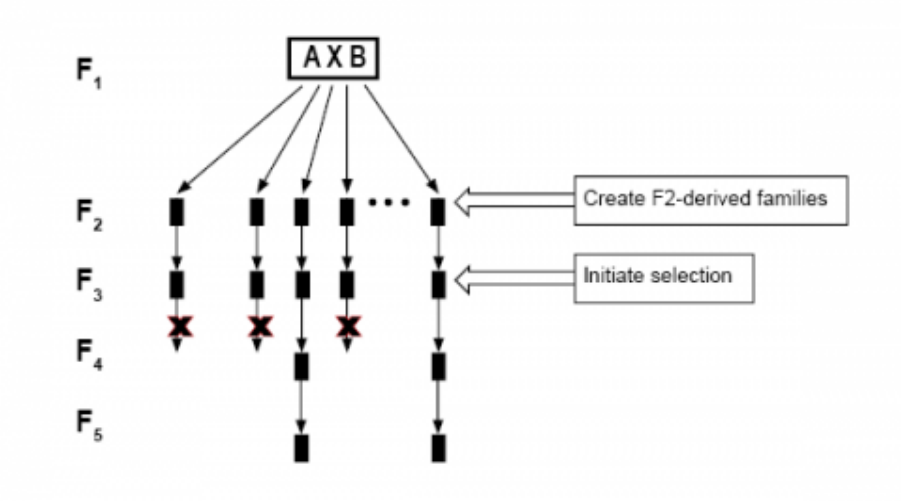


Fig. 11 Selection during inbreeding in a self-pollinated crop.

Furthermore, discarding poor-performing progeny early saves testing resources. However, seed supplies may be limited in early generations. Moreover, the use of homozygous lines eliminates “noise” in the data due to segregation and increases selection accuracy.

The trade-off is between producing more accurate estimates of progeny means vs. sampling a larger number of progeny per cross.

Number of Progeny

Consider the number of progeny to produce and test.

What number of progeny must be tested to assure that at least one truly superior progeny will be identified if, in fact, it exists in that population?

General Guidelines for Determining the Number of Progeny

The number of progeny produced per breeding cross depends on trait heritability as well as the expected proportion of truly outstanding progeny in the specific population and the selection intensity applied.

- There should be at least 200 progeny per population for trait of low to medium heritability.
- If genetic diversity is limited (as it is with soybean), you need even more progeny to find unique gene combinations. More than the minimum of 200 progeny per population is needed.
- Note that the commercial soybean program example suggests population sizes upwards of 500 (200-500 are selected after evaluation for some unspecified characteristics, perhaps disease screening or morphology).
- You must consider the number of traits to be improved and the intended selection intensity for each.

Breeding Methods

How will progeny be advanced to produce materials for testing? This ultimately relates to choice of *breeding method*. Testing will take place in various stages through the process pipeline. Selections made as a result of one stage of testing will be advanced to the next stage, which often involves the next generation. The breeding method is referred to as the "recombination unit."

Breeding methods involve the progression of selected materials through the breeding pipeline. Since advanced stages of testing are intended to provide a view of performance of entries as released cultivars, the breeding method(s) utilized are also transitioning selected lines to the genetic state required for release as a new, improved cultivar.

Notice the example of a commercial soybean breeding program in Chapter 1 which called for a "modified single-seed descent" breeding method to advance S_0 plants to S_1. [Example: Commercial Soybean Improvement Program in Chapter 1].

How is this method employed?

The Pedigree Breeding Method

Single-seed descent is a variation on the pedigree breeding method (Fig. 12), which is a method used in the inbreeding of populations of self- or cross-pollinated species with the goal of developing inbred lines.

The pedigree breeding method typically starts with an F_2 population, with self-pollination of each generation until fully inbred lines are generated. Selections can be made in each generation as desired, advancing best **families**, best rows within a family, or best plants within rows. Thus, the pedigree breeding method can be used to select among families as well as within families.

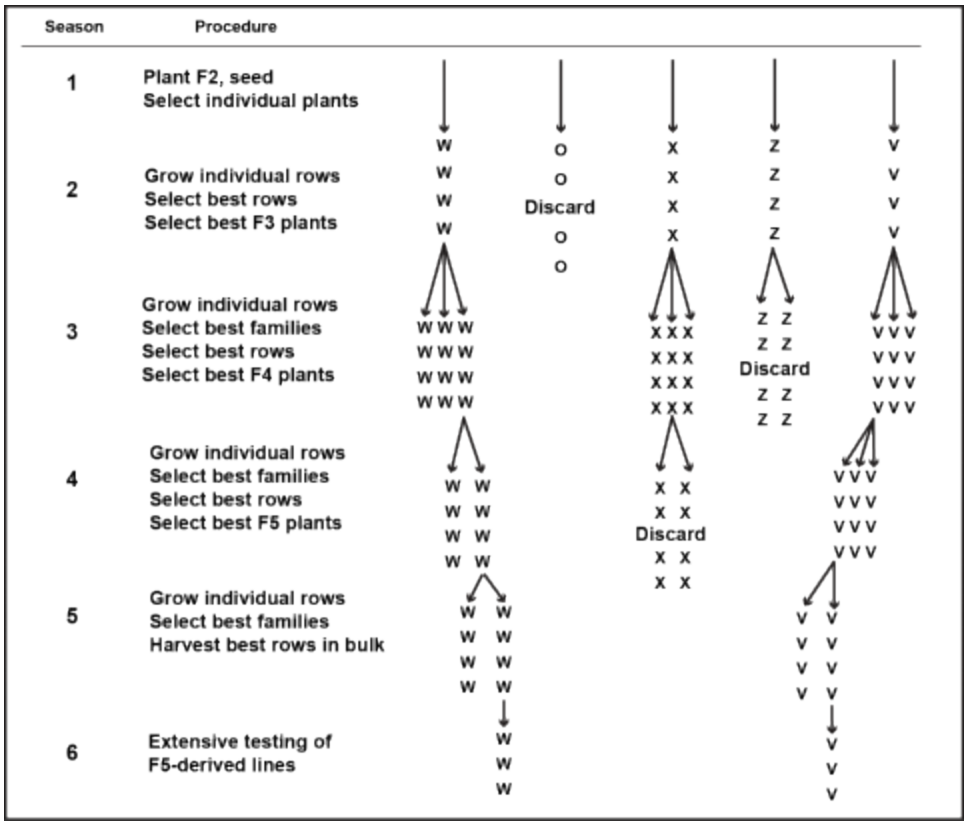


Fig. 12 Illustration of the pedigree breeding method from Fehr. 1978.

Single-Seed Descent: A Version of the Pedigree Method

Single-seed descent involves the advancement of F_2 lines through self-pollination to homozygosity. This method involves harvesting a single seed from each plant at each generation, bulking the individual seeds, and planting out the entire bulk to represent the next generation. Testing is initiated once the desired level of homozygosity is reached.

This method is an easy way to maintain large populations through the inbreeding process and can be implemented in a relatively short timeframe using off-season nurseries and greenhouses.

Modified Single-Seed Descent

The modified single-seed descent method, referenced in the example of a commercial soybean breeding program, calls for retaining a single pod containing two to three seeds from each F_2 plant (instead of a single seed) and bulking across the population. This modified method serves to expand the number of $F_{2:3}$ families represented in the breeding population.

Evaluation and Selection

Evaluation and Selection

The *cycle* of cultivar improvement and, thus, the *process* of cultivar development (Fig. 1) embodied by the process pipeline, centers on the product target. Genetic variation represented in the breeding populations is exploited to make genetic gain toward the defined product target.

Having decided what materials to evaluate, the next key decisions involve:

- the basis for evaluation.
- how the evaluations are conducted, and
- criteria for selection.

The key is to focus on these to *maximize* the ability and probability of identifying *outstanding* individuals that meet or exceed the product target if these are present in the breeding populations!

Trait Screens

Screens are needed to measure traits of interest.

$$\frac{t}{ha} = \frac{\text{shelled seed weight (in kg)}}{plot} \times \frac{100 - \%\text{mst}}{87} \times \frac{1t}{1000kg} \times \frac{\text{number of plots}}{ha}$$

With soybean, seed yield is adjusted to a standard 13 percent grain moisture and expressed in weight per land unit.

Fig. 13 Threshing soybean plants. Photo courtesy International Institute of Tropical Agriculture, Ibadan, Nigeria.

Fig. 14 Harvesting soybean with a single-row plot combine harvester. Photo courtesy of Almaco.

For example, seed yield in soybean is typically measured using a seed thresher (Fig. 13, above) or a combine (Fig. 14) based on weight of the seed per plot.

Fig. 15 Grain Moisture Meter LDS-1G For Agriculture/ food Photo from Alibaba.com.

Seed moisture may be measured using a moisture meter (Fig. 15). Typically, the moisture meter is incorporated in the combine to simplify data collection.

The Need for Protocols

Protocols are needed to ensure that trait measurements are performed accurately, uniformly across the entire experiment, and in conformity with accepted practices.

Note that with the soybean product target example, there is a need for a protocol to describe in detail the exact procedure to use in evaluating pod shattering (also called pod dehiscence).

Romkaew and Umezaki (2006) confirmed that the oven-dried method is a reliable way to discriminate between cultivars that are resistant and susceptible to pod shattering.

METHOD: 30 pods, each with 2 seeds, which have been maintained at <50% relative humidity, are exposed to 60°C in an oven for 7 hours. The number of opened pods is counted and the percentage of pod shattering is computed.

Crop Ontology

Wait, there's more…

The performance measurement must be expressed in terms of a scale known to the Community of Practice

for the particular crop. Uniform nomenclature is important to sharing information and results, whether internal to the breeder's organization or to the plant breeding community at large.

The BMS (Breeding Management System) features an ontology for a number of crops.

Choice of Testing Sites

What is the relationship between the testing locations and the market region?

Testing locations are intended as "samples" of the market region. Therefore, it is essential that all the elements of the production areas in the market region are represented and reflected in the characteristics of the testing locations:

- Geographical position.
- Season (if there is more than one per year).
- Soil type.
- Planting dates.
- Cultural practices of farmers in the region: tillage, fertilizer regimes, harvesting methods, irrigation, microbial inoculation of the soil, crop rotation, etc.

Ideal testing locations represent the market region well, and they also enable discrimination among test entries (i.e. differences among test entries are apparent). The uniformity (lack of variation) within a test site allows for the variation in performance among test entries to be front-and-center.

Field Uniformity

For testing sites, choose fields that are uniform (Fig. 16).

What factors contribute to uniformity?

- Topography.
- Soil type.
- Previous crop.
- Planting preparation.
- Access to moisture.
- Protection from disturbance (if goats get into the field, they probably won't sample evenly across the field!).

Fig. 16 Plowed field near Ames, IA. Iowa State University extension plowed field images

Any differences within a field should be associated with different blocks to the extent possible. For example, confine to one block a low spot in the field that may harbor standing water. That way, the block (and not the whole location) can be discarded if necessary.

Block sizes that are too large result in a lack of uniformity and homogeneity within. The choice of the proper experimental design safeguards the homogeneity of the block. For example, advanced yield trials often include many entries grown in larger-sized plots. Because one block would be too large to maintain homogeneity, an incomplete block design is typically employed to accommodate all and effectively partition any environmental effects.

Appropriate Experimental Design for Trials

Conduct evaluations using appropriate experimental designs to partition variation due to genotype, the environment, and other sources.

Proper execution includes:

- Randomization to guard against bias.
- Replication to capture natural variation among **experimental units** to be treated alike.
- Use of blocking to partition variation due to environmental influences.

Best Practices in Data Collection

Data integrity must be safeguarded in data collection!

- Prepare to collect data before going to the field or greenhouse. This may mean printing forms for data collection that reflect the field layout. (You do not want to search for the plot number on a form when recording your observed data value.)

- Staff to properly collect data effectively. Having one person to evaluate and another person to write in the observed score on the form can work well. It is best to have the same person rate the entire trial; however, if that is not feasible, assign those ratings to individual reps.
- Data collection programs like Field Book (Fig. 17) may be useful, facilitating data collection on a Kindle or an Android phone.

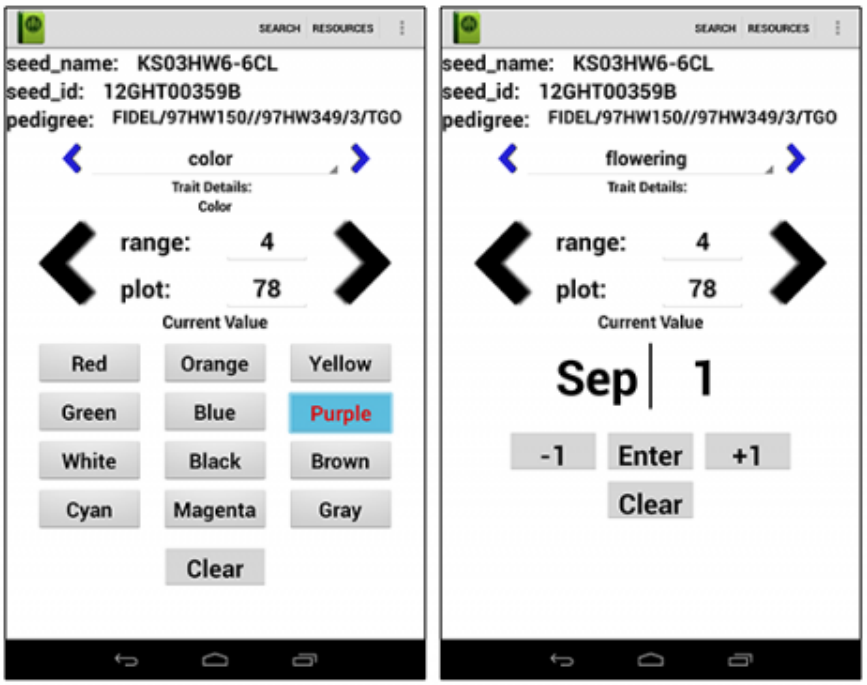


Fig. 17 Field Book allows input screen customization for categorical (left) and date trait (right) formats, facilitating easier use and limiting entry mistakes. Image from Rife and Poland, 2015.

Traits to Observe and Record

What traits need to be observed and recorded?

Certainly include all the traits specified in the product target. Sometimes, it can be useful to collect data for additional traits, especially those that correlate with target traits.

For example, soybean seed yield is **positively correlated** with:

- Number of pods.
- Number of seeds per pod.
- Fruiting period.
- 100-seed weight.

Recording correlated traits provides an opportunity to collect additional information on key traits.

Furthermore, sometimes it is much more feasible to measure a highly correlated trait that may be easier or less expensive to collect, or that can be collected earlier in the crop life cycle, shortening the breeding cycle.

Indirect Selection

Evaluation and selection based on a correlated trait is an example of **indirect selection**.

To improve Character X, we might select for another Character Y. In this case, Character Y is considered a "secondary trait."

What is the expected response to selection in Character X when selection is applied to Character Y?

The expected correlated response of X when selecting based on Y is:

$$CR_X = ih_X h_Y r_{G_{Xy}} \sigma_{P_X}$$

Given that the response of Character X selected directly is:

$$R_X = ih_X \sigma_{A_X}$$

Indirect selection will result in relatively greater genetic gain for Character X than direct selection, if:

$$r_{G_{Xy}} h_Y > h_X$$

Examples of Secondary Traits

Some examples of secondary traits used in crop improvement include:

- Canopy temperature as a proxy for drought stress tolerance (Furbank and Tester, 2011) (Fig. 18).
- Root mass under water-limited conditions as a proxy for drought stress tolerance.
- Leaf senescence below the ear as a proxy for low nitrogen tolerance in maize (Banziger and Lafitte, 1997).
- Yield component characteristic (e.g., 100-kernel weight) as a proxy for grain yield.

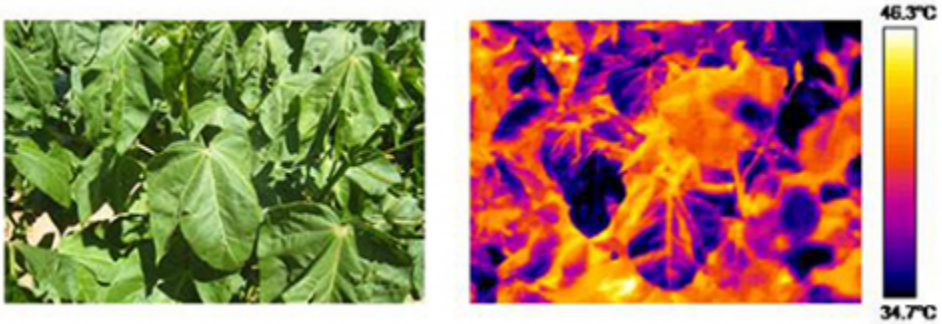


Fig. 18 Photograph and FLIR (Forward Looking Infrared) image of water-limited cotton plant captures the variation in leaf temperature. From Carmo-Silva et al. 2012.

Choosing Secondary Traits

In choosing secondary traits, consider those that:

- Correlate genetically with the trait of interest in the given target market environment.
- Have higher heritability than the trait of interest (less non-additive genetic variation), preferably above 0.6.
- Exhibit genetic variation.
- Are not associated with poor performance in non-stressed environments.
- Easily, cheaply measurable (need a good screen!).
- Can be measured on a per-plant or per-plot basis, preferably in a non-destructive manner.

Ancillary Traits

There may be other instances where you wish to collect traits other than those in your product target.

1. If you wish to estimate harvest index, calculations involve the “seed yield to biomass” ratio. Therefore, biomass must be measured.
2. Other characteristics may be collected to gather and present information in order to categorize potential new cultivars. For instance, soybean characteristics typically collected to help position any prospective new cultivars appropriately in the marketplace include:

- Presence of pubescence.
- Seed coat color (16 categories).
- Hilum color.
- Growth type: determinate or indeterminate.

Thus, you may have good reason to expend the time, effort, and resources to measure, record, and analyze other traits, besides those in the product target, to help identify outstanding progeny that may be prospective improved cultivars.

Consider all your data needs before implementing the trial.

Multiple Trait Selection

Product targets are never about a single trait. Typically, plant breeders seek to develop cultivars improved for a number of traits. Let us look at different ways we can approach selection considering that our product target involves multiple traits.

Multiple-trait selection approaches include:

- Tandem selection.
- Independent culling levels.

- Index selection.

Tandem Selection

Tandem selection is a methodology involving selection for one trait at a time (Fig. 19). Selection for each trait is sequential; selection thresholds and selection intensities for each trait are independent. The shaded area indicated the selected individuals in the population.

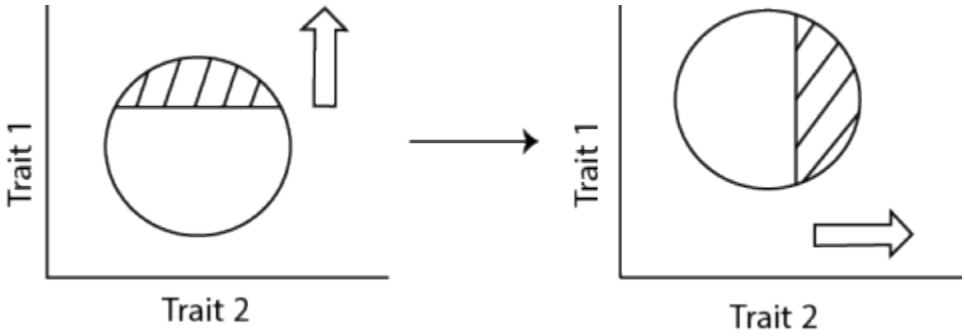


Fig. 19 Selection for one trait at a time. Adapted from Bernardo, 2010.

Tandem selection is a common practice. It is often deployed for traits that affect adaptation or involve stress tolerances.

Examples:

- Tropical maize populations which may be used for improving temperate maize have been selected for photoperiod sensitivity prior to selection for traits such as yield.
- In elite materials, selection for disease resistance may precede selection for yield.

Independent Culling Levels

Independent culling is a method involving selection for more than one trait in a single step (Fig. 20). Thresholds are established for each trait and only those individuals that meet all trait thresholds are selected. With this method, individuals can be culled on the basis of a single trait without having to wait for or even collect observations of the other trait(s).

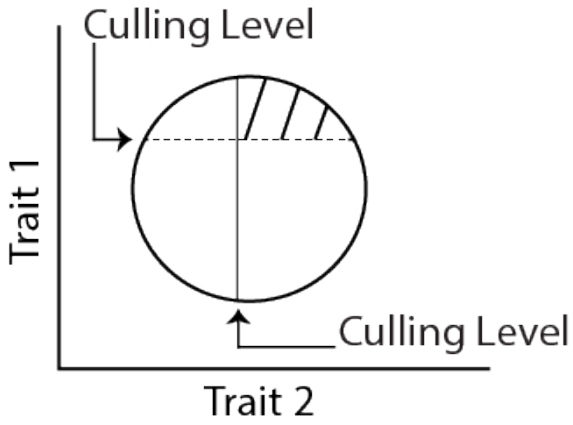


Fig. 20 Selection for more than one trait at a time. Adapted from Bernardo, 2010.

Index Selection

Index selection is a method involving selection for more than one trait simultaneously on the basis of a single index value (Fig. 21). It accounts for the relative superiority of individuals for all of the traits included in the index.

The selection index is usually a linear function of the various traits involved. Each trait is weighted according to its importance. Furthermore, the index typically takes into account the genetic correlation between pairs of traits so that progress can be made even with negatively correlated traits.

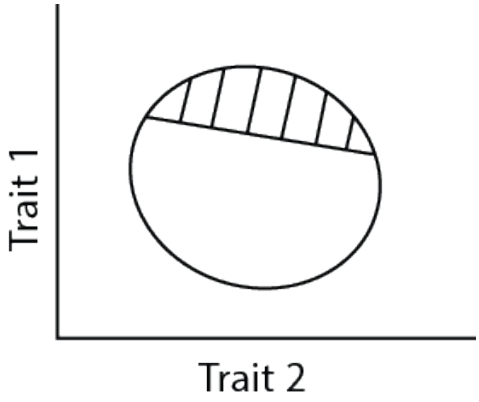


Fig. 21 Index selection for traits that are negatively correlated traits. Adapted from Bernardo, 2010.

Smith-Hazel Index

Index selection is a form of indirect selection. As for selection indices, there are a number of types. All are based on the original index independently developed by Smith (1936) and Hazel (1943).

The Smith-Hazel Index takes the form:

$$I = b_1 X_1 + b_2 X_2 + \ldots + b_n X_n$$

$$I = \sum b_i X_i$$

where:
b_i is the weight of trait i,
X_i is the phenotypic value for trait i.

i is calculated for each individual in the population. The index value becomes the new variable upon which breeding progeny are evaluated. Individuals with the highest values are selected.

Weights in the Smith-Hazel Index

How are weights calculated when using the Smith-Hazel index?

Weights take into account the phenotypic and genetic covariances among the traits, as well as the economic value of each trait. Weights can be calculated by solving the following equation:

$$b = P^{-1} G_a$$

where:
b is an $nx1$ vector of the b_i values,
P^{-1} is the inverse of the nxn matrix of phenotypic covariances among the traits,
G is thenxn matrix of genetic covariances among the traits,
a is an $nx1$ vector of economic weights for the traits.

Note that G and a may be difficult to estimate.

See Quantitative Genetics: Multiple Trait Selection *(coming soon!)* for more on index selection.

Efficiency in Selecting for Multiple Traits

Which multiple trait selection approach is most effective?

According to Hazel and Lush (1942) the order is:

Selection index > Independent culling > Tandem selection

Testing Regime

To facilitate selection, the pipeline process embodies a testing regime to evaluate progeny, resulting from the various breeding crosses, for all the traits of interest.

Let's review the testing regime outlined in the Commercial Soybean Improvement Program example that was introduced in Chapter 1:

Table 5 New Line Development and New Line Evaluation in a commercial breeding program for soybean. Adapted from Bernardo, 2010.

Season[1]	Activity
Winter 0	Make breeding crosses
Summer 1	Self or BC each population
Winter 1a[2]	Grow 200 F_2 or BC_1 populations (i.e. S_0 generation) that have been formed in previous years Advance the S_0 plants to the S_1 generation by a modified single-seed-descent method, retaining single pod (instead of a single seed) with 2-3 seeds; bulk by population
Winter 1b	Plant S_1 seed bulk; Select 200-500 plants (~350 per population) and save selfed (i.e. S_2) seed
Summer 2	Yield trials for 70,000 S_2 families (across all populations) in unreplicated trials at 1-2 locations; Select the best 5,000 based on yield performance; Save S_3 seed from the trials
Summer 3	Yield trials for 5,000 S_3 families at 3-5 locations; Select the best 200 based on yield performance; Save S_4 seed from the trials
Summer 4	Yield trials for 200 S_4 families at 15-25 locations; Select the best based on yield performance; code selected as 'experimental' lines; Save S_5 seed from the trials
Winter 4	Increase seed of experimental lines
Summer 5	Yield trials of experimental lines at 20-40 locations On-farm strip tests (i.e.150-300 m^2 plots) at 20-100 locations
Summer 6	Yield trials of 'advanced' lines at 20-50 locations On-farm strip tests (i.e. 150-300 m^2 plots) at 30-500 locations
Fall 6	Release 0-5 new varieties

[1]Summer represents the main growing season, winter denotes off-season activities; number after season indicates the year in the development pipeline

[2]Winter nurseries may be grown back-to-back in the same winter season

Aspects of the Testing Regime

As noted earlier, this example serves a market region with one growing season, which is referred to as "Summer." "Winter" designates off-season activities.

We observe the following with respect to the testing regime:

- The main focus of testing is grain yield in this example (no other trials other than yield trials are mentioned, although selection among individual S_1 plants is indicated). High grain yield is obviously the trait of the utmost priority in producing new varieties (in this case).
- A large number of breeding populations are created.
- Yield trials are initiated with a very large number of S_2 families and continue every growing season in the market region.
- Selection intensities are high (only a small proportion of the tested families are advanced to the next generation of testing).
- The number of locations for yield testing increases as the number of tested families decreases.
- Lines that are selected after Summer 4 yield trials (third year of yield testing) are **coded** as "experimental" and progressed to wide-area testing the following growing season. Lines that continue to meet or exceed product targets are progressed to "advanced" line status in further wide-area testing.
- Wide-area testing includes "on-farm trials" managed by potential seed customers. Wide-area testing could also include National Performance Trials required for the registration of a new variety.
- Overall, there are at least five years of comprehensive performance testing: five seasons of "research" yield trials, and two seasons of wide-area testing which are initiated after three seasons of research yield trials.
- Because soybean is a predominantly self-pollinated crop (pollination typically takes place before flowers open), seed for planting the following year's trials can be retrieved from yield plots (after recording yield measurements). Seed production for wide-area trials that require significant seed amounts is done during the off-season to save time.

Testing Progression

Let us examine the progression of testing in the example program.

Preliminary yield trials are conducted in Summers 2 and 3. Preliminary yield trials are characterized by a large number of entries observed at a small number of locations.

In the example program, there is a third year of yield testing with only a fraction of the number of lines originally created (i.e. 200 S_4 families).

Trials are unreplicated, meaning that there's only one rep of the test at each location.

The testing regime takes into account the various factors contributing to phenotypic variation so that variation associated with "genotype" can be partitioned and differences between entries can be perceived.

We know that:

$$V_P = V_G + V_E + V_G x E + error$$

$$and$$

$$V_E = V_n + V_r + V_L + V_Y$$

where:
V_P is phenotypic variation
V_G is genetic variation
V_E is environmental variation
V_{GxE} is variation due to the interaction between genotype and environment
V_n is variation among plants in an experimental unit (e.g. plot)
V_r is variation among replications of experimental units
V_L is variation among locations
V_Y is variation among years.

These formulas indicate that, to elucidate variation due to genotypic variation, variation due to environment must be accounted for and error must be well controlled.

To accomplish this, the example demonstrates the following:

- Experimental units containing more than one plant are employed, especially in measuring a low- to medium-heritability trait like grain yield.
- Multiple locations are needed to sample an increasing range of the target market as testing progresses.
- Multiple years (i.e. seasons) of testing are needed to ensure an adequate sample of climatic conditions in the market region.

The regime in the example program also allows for an assessment of GxE (Genotype by Environment interaction). Ideally, a new improved cultivar will perform *best* or *among the best* across the entire market region.

Unreplicated Yield Trials

Why does the example program advocate *unreplicated* yield trials? Isn't replication needed to identify V_r (i.e., variation among experimental units treated the same)?

Not necessarily. Work by Melchinger et al. (2005) showed that, for a given number of observations of each entry, maximum gains from selection are facilitated by maximizing the number of locations at the cost of replication within locations (Fig. 22). In other words, single-rep trials at more locations provide the greatest opportunity for gains in grain yield for any case where there are two or more locations involved. Note that two locations with one rep each facilitate more genetic gain than selection based on one location with two replications.

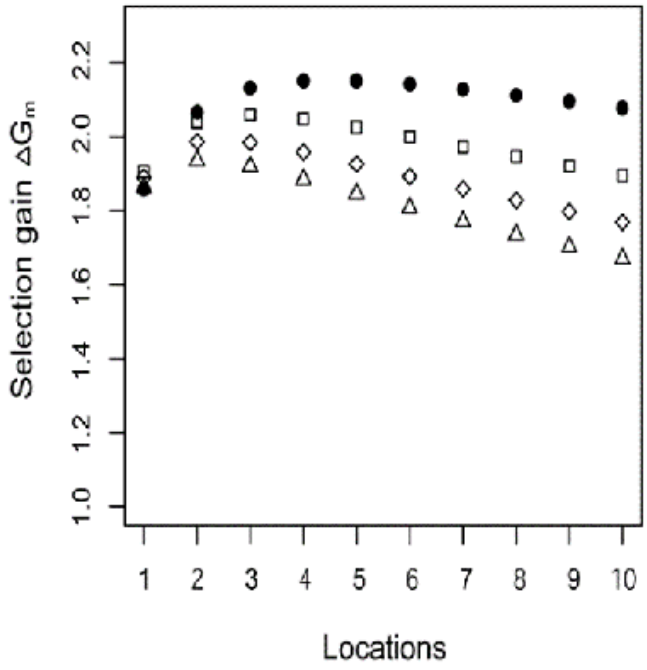


Fig. 22 Maximum selection gain versus number of locations and level of replication. From Melchinger et al. 2005.

Symbols indicate the number of replications: $(r) : r = 1(\bullet), r = 2(\square), r = 3(\lozenge), r = 4(\triangle)$.

Early Generation Yield Testing

In Summer 2, there are 70,000 S_2 families grown at 1-2 locations. Only the best 7% will be advanced to the next yield trial.

Preliminary Yield Testing

What is the genetic state of these testing materials? Are these lines fully inbred, or is there some level of heterozygosity present?

Click here to reveal the answer

The S2 lines are F4 equivalents and, therefore, are not fully homozygous. On average, they are 87.5% homozygous.

This is an instance of early-generation testing.

Experimental Designs Employed

Overall, there are at least five years of comprehensive performance testing: five seasons of “research” yield trials, and two seasons of wide-area testing which are initiated.

What experimental designs would be suitable for the various stages of yield testing?

Research trials feature:

- 70,000 entries in Summer 2
- 5,000 entries in Summer 3
- 200 entries in Summer 4

In Summers 5 and 6, research trials continue. Only a small number of entries are still being advanced from previous trials. However, the type of testing done in Summers 5 and 6 is very comprehensive coverage of the market area. This testing is meant to provide a good sampling of the types of environments that would be encountered should one or more of these lines be released as a new variety. These tests would include other entries besides the advanced selected lines from the particular breeding crosses represented in Winter 1a. These tests may include a broad set of current cultivars, competitive cultivars, advanced lines from previous years, etc. Therefore, these tests are likely to include more than 30 entries overall.

Suitable Experimental Designs

Suitable experimental designs are those that facilitate accurate and precise estimates of mean performance, even with a large number of entries.

Incomplete block designs are suitable because they partition a replication of the full experiment into smaller, more homogeneous units of land. (Note: with a large number of entries, one full set of entries is likely to occupy a large space in a field, making it difficult to maintain homogeneity over this large area.)

You may be familiar with these types of incomplete block designs:

- Lattice
- Alpha-lattice
- Row-column design

These are commonly used in yield testing when trials include a large number of entries.

Furthermore, within the context of these types of experimental designs, augmented designs can be considered as a means to increase precision with unreplicated trials.

Incomplete Designs

Would incomplete block designs also be suitable for conducting the on-farm trials? Why or why not?

Click here to reveal the answer
Yes, incomplete block designs are suitable for farmer-participatory trials, although with increased plot sizes, block sizes would be smaller.

Augmented designs would be a good fit as well.

Number of Locations

In the example program, the number of **locations** (i.e. environments) expands each year of testing… and a range for the number of locations at each point of progression in testing is given.

What criteria factor into the determination of the number of locations?

In the example, Summers 2 and 3 are winnowing down the number of overall entries to a manageable number, which will be screened in Summer 4 to assess the magnitude of yield differences among the lines and the lines compared to **check** entries. The number of locations is ramped up as the number of entries is pared down. Budgets are an important consideration in determining the number of locations for these seasons.

Summer 4 represents the critical assessment of the remaining lines for grain yield. The Summer 4 yield trial is critical to identifying lines that meet the stated selection benchmarks. Therefore, the number of locations to be utilized in Summer 4 is pivotal to ultimately determining success in achieving the product target.

Determining the Number of Locations Needed

How many locations are needed to detect a particular Least Significant Difference (LSD)?

The breeder sets the desired probability level (e.g., α = 0.05) and can plug in estimates of the error variance (V_{error}) and the variance associated with genotype by environment interaction (V_{GE}) obtained in previous trials.

For a two-tailed test with α = 0.05:

$$LSD_{0.05} = t_{\alpha/2}\sqrt{2(\frac{V_{error}}{re} + \frac{V_{GE}}{e})}$$

where:

r = number of replications per location,
e = number of locations.

When the number of entries in the trial is large (i.e. >30), the critical value of $t_{\alpha/2}$ is about 2.0.
Thus, the equation becomes:

$$LSD_{0.05} = 2\sqrt{2(\frac{V_{error}}{re} + \frac{V_{GE}}{e})}$$

The equation can be rearranged to solve for the number of locations required (i.e., e required) to facilitate detection of a certain LSD at the α = 0.05 significance level:

$$e_{required} = (\frac{8}{(LSD_{0.05})^2}(\frac{V_{error}}{r} + V_{GE})$$

If the breeder wishes to ascertain a 0.3 t/ha increase in yield, then the equation becomes:

$$e_{required} = (\frac{8}{(0.3)^2}(\frac{V_{error}}{r} + V_{GE})$$

With unreplicated trials, $r = 1$, and the equation can be solved with estimates of the error variance V_{error} and the variance associated with genotype by environment interaction V_{GE} obtained in previous similar trials.

$$e_{required} = \frac{8}{0.09}(V_{error} + V_{GE})$$

Furthermore, the number of locations can also be determined in the case of replicated trials by plugging in the appropriate value of r.

Estimates of Variance Due to Error and to GxE

Note that estimates of error variance (V_{error}) and the variance associated with genotype by environment interaction (V_{GE}) are easily obtained based on the Expected Mean Square (EMS) terms for the sources of variation (SOV) in the ANOVA. Previous trials of similar size, market area, and trait of interest may be used.

In the following example (*n*=genotypes, *e*=environments, and *r*=replications), genotypes and environments are considered random; the EMS terms are given for pertinent SOV.

Table 6 Sources of variation and expected mean squares.

SOV	DF	MS	EMS
Environments (E)	e-1	n/a	n/a
Blocks/E	e(r-1)	n/a	n/a
Genotypes (N)	n-1	$MS_{genotype}$	$V_{error} + rV_{GE} + reV_{genotype}$
GxE	(n-1)(e-1)	MS_{GE}	$V_{error} + rV_{GE}$
Error	(n-1)(r-1)e	MS_{error}	V_{error}
Total	nre-1	n/a	n/a

An estimate of V_{GE} can be easily calculated using the mean square for GxE and the mean square error:

$$V_{GE} = \frac{MS_{GE} - MS_{error}}{r}$$

Patterns of GxE

Let's explore further into GxE…

There are three main patterns of GxE exhibited (Fig. 23). Considering performance (*y* axis) of two cultivars (blue and green lines) at two environments (*x* axis), GxE pattern may look like this:

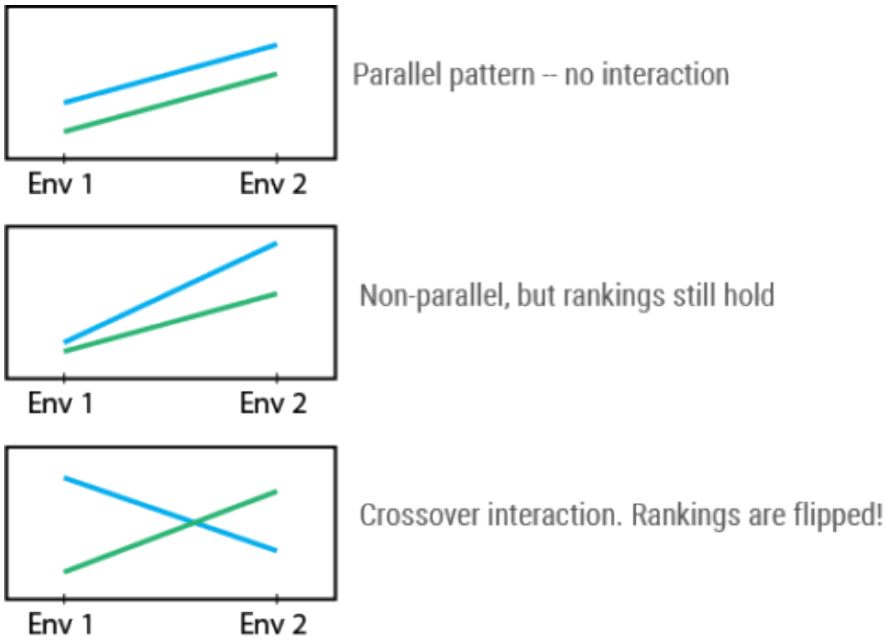


Fig. 23 Patterns of GxE: parallel, non-parallel, and crossover interaction.

With crossover interaction, the ranking of genotypes is flipped between the two environments. That is, blue is "best" in the first environment whereas green is "best" in the second environment. This makes selection

complicated! After all, as plant breeders, we are interested in identifying the "best" genotype for our entire target market if possible.

An ANOVA may indicate the presence of GxE through a statistically significant F-test; however, this does not indicate whether rankings are affected. Further probing is required (e.g., mean comparisons of entries by environment).

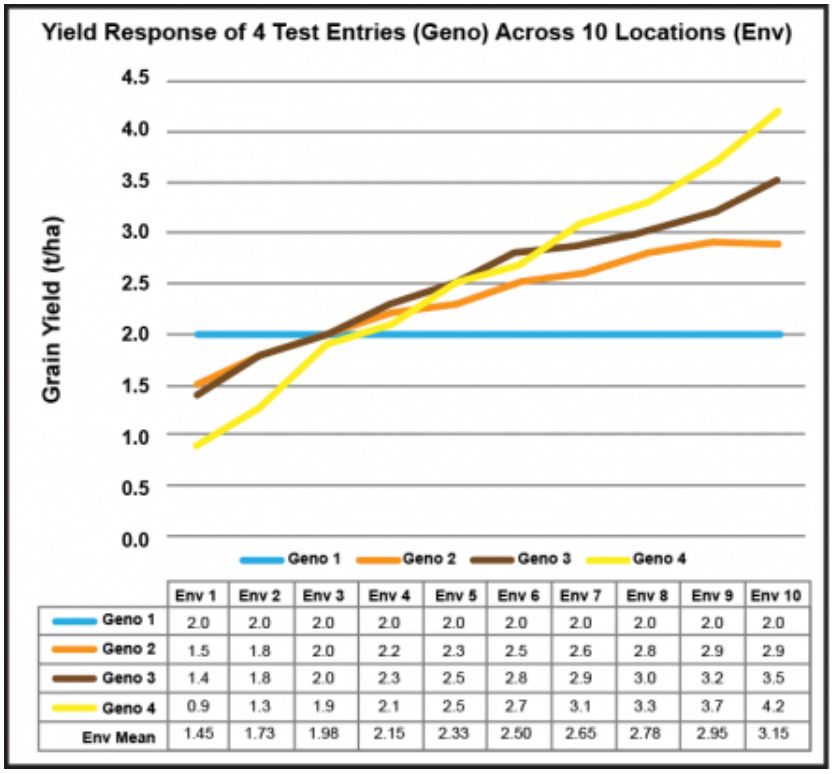
Managing GxE

As a plant breeder, there are a number of ways to deal with GxE:

- Ignore it
  - Testing done in wide range of environments and selections based on mean performance.
  - Does not recognize best cultivars for specific environments.
- Reduce it
  - Partition the population of environments into smaller, more homogeneous regions.
  - Make selections by region.
- Exploit it
  - Reduce GxE by partitioning environments into mega-environments.
  - Aims to identify cultivars best suited to specific environments or subsets of environments.
  - Also considers "stability" across environments.

What is "Stability"?

The concept of "stability" involves the performance of a line relative to other lines across a range of environments that represent the target market (Fig. 24). If environments are ranked in terms of their effect (mean performance of all entries) and charted from low to high, the response of the lines tested can be graphed. For example, four soybean lines (i.e. genotypes) are ranked as to their performance for seed yield at each of 10 locations (i.e., environments):



	Env 1	Env 2	Env 3	Env 4	Env 5	Env 6	Env 7	Env 8	Env 9	Env 10
Geno 1	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Geno 2	1.5	1.8	2.0	2.2	2.3	2.5	2.6	2.8	2.9	2.9
Geno 3	1.4	1.8	2.0	2.3	2.5	2.8	2.9	3.0	3.2	3.5
Geno 4	0.9	1.3	1.9	2.1	2.5	2.7	3.1	3.3	3.7	4.2
Env Mean	1.45	1.73	1.98	2.15	2.33	2.50	2.65	2.78	2.95	3.15

Fig. 24 Performance of four genotypes across environments.

Mean Performance Across Environments as a Measure of Stability

Regression analysis can be implemented to calculate the linear relationship between mean performance and environments for each genotype (b_i).

We can examine mean performance as a measure of yield stability (Fig. 25):

- Genotype 1 is constant in its performance across all 10 environments. In this case, $b_i = 0$. Moreover, this entry yielded more than the other lines under less favorable growing conditions.
- Genotype 3 displayed a response similar to the slope of the mean response of all four genotypes across environments (regression not shown).
- Genotype 4 responded best to more favorable environments, displaying its genetic potential. However, it yielded less than the other lines under stressful conditions.

Which genotype exhibits the greatest yield stability?

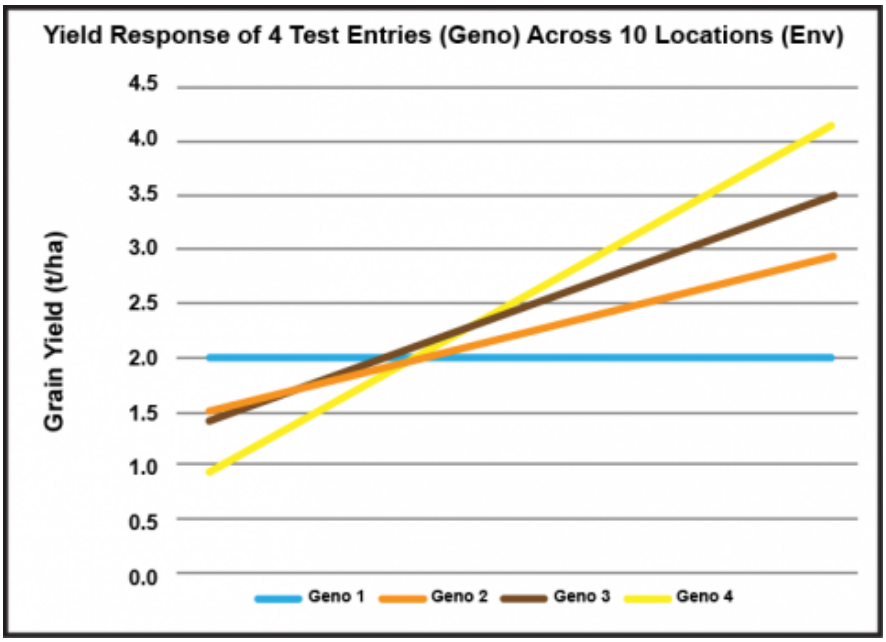


Fig. 25 Performance expressed as the linear relationship across environments.

Deviations From Regression as a Measure of Stability

Although we have examined stability in terms of response across environments, another measure of stability of a genotype can be considered: deviations from the fitted regression line. Excessive deviations from the regression line would suggest erratic response to the range of environments tested.

For more detail on stability analysis, see Bernardo 2010, Chapter 8.5.

Partitioning Environments to Deal with GxE

Another approach to managing/exploiting GxE is to partition environments into homogeneous subgroups or **mega-environments**. In essence, this leads to the formulation of multiple product targets, one for each mega-environment. Then, testing is conducted accordingly.

Partitioning environments can be facilitated through cluster analysis (the same type of analysis used to partition genetic diversity; (see Choosing Parents) based on data from performance trials across locations and years. Bernardo (2010) offers suggestions on ways to compute distance estimates between pairs of environments.

In Chapter 6 of this course, we will examine a method called GGE Biplot to determine whether a target market region comprises more than one mega-environment. GGE Biplot can also be utilized to assess the stability of tested genotypes.

Advanced Testing

Advanced testing (Summers 5 and 6) focuses on lines that have met or exceeded specifications of the product target in the earlier years of testing (i.e. coded lines). Advanced testing is characterized by:

- Maximal numbers of research testing sites.
- Farmer participation in testing, with larger plot sizes employed to provide a better view of overall performance under real-life field conditions in the market region.
- Wider exposure of farmers to potential new cultivars by maximizing the number of farmer-participatory sites. More farmers have the opportunity to gain experience and familiarization with the potential new products.

The purpose of advanced testing is to confirm coded lines meet or exceed all the specifications in the product target, across an even wider range of samples from the population of environments representing the target market. Furthermore, this testing contributes to the assembly of information for potential product release to farmers for cultivation, and for varietal registration.

Ideally, there is at least one coded line that represents a new "best" for the target market.

Identifying Potential New Products

Potential new cultivars will meet or exceed all the specifications in the product target throughout the entire testing regime. Additionally, such candidate lines will not exhibit any negative characteristics that would make them undesirable to farmers or other stakeholders in the value chain (Fig. 25). Those lines with the best performance overall will be considered for possible release as new, improved cultivars.

According to the soybean breeding pipeline example, zero to five new cultivars may come through as candidates for release. As demonstrated in this chapter, all aspects of the product pipeline are interrelated and require integration for success and value.

Fig. 25 Improved soybean varieties and production tactics demonstrated to smallholder farmers in Wa, Ghana; photo courtesy of USAID Feed the Future Soybean Innovation Lab.

If the product pipeline is designed well, integrated fully, and implemented effectively, there is a higher probability that the number of potential new products will be in the 1-5 range. *We aim for success!*

References

Banziger, M., H.R. Lafitte. 1997. Efficiency of secondary traits for improving maize for low-nitrogen target environments. Crop Sci. 37:1110-1117.

Bernardo, R. 2010. Breeding for Quantitative Traits in Plants. 2nd ed. Stemma Press, Woodbury, MN.

Carmo-Silva, A.E., M.A. Gore, P. Andrade-Sanchez, A.N. French, D.J. Hunsaker, M.E. Salvucci. 2012. Decreased CO2 availability and inactivation of Rubisco limit photosynthesis in cotton plants under heat and drought stress in the field. Environmental and Experimental Botany 83: 1-11. dx.doi.org/10.1016/j.envexpbot.2012.04.001

Cockerham C.C. 1963. Estimation of genetic variances. p. 53- 94. In statistical genetics and plant breeding. National Academy of sciences, NRC.

Falconer, D.S. 1989. Introduction to Quantitative Genetics, 3rd ed. Longman Scientific & Technical, Essex, England.

Fehr, W.R. 1978. Breeding. In A.G. Norman (ed.), Soybean Physiology, Agronomy, and Utilization, p120-155. Academic Press, New York.

Furbank, R.T., M. Tester. 2011. Phenomics – technologies to relieve the phenotyping bottleneck. Trends in Plant Science 16(12): 635-644.

Harlan, J.R., J.M.J. de Wet. 1971. Toward a rational classification of cultivated plants. Taxon 20(4): 509-517.

Hazel, L.N., J.L. Lush. 1942. The efficiency of three methods of selection. J. Heredity 33: 393-399.

Heffner, E.L., M.E. Sorrells, J.L. Jannick. 2009. Genomic selection for crop improvement. Crop Sci. 49:1-12.

Irwin, S.V., R.V. Kesseli, W. Waycott, E.J. Ryder, J.J. Cho, R.W. Michelmore. 1999. Identification of PCR-based markers flanking the recessive LMV resistance gene mo1 in an intraspecific cross in lettuce. GENOME, 42:982-986.

Jacob, C., B. Carrasco, A. R. Schwember. 2016. Advances in breeding and biotechnology of legume crops. Plant Cell Tissue and Organ Culture 127:561-584. DOI 10.1007/s11240-016-1106-2

Melchinger, A.E., C.F. Longin, H.F. Utz, J.C. Reif. 2005. Hybrid maize breeding with doubled haploid lines: Quantitative genetic and selection theory for optimum allocation of resources. Pp8-21. In Proc. 41st Annual Corn Breeder's School, Mar 2005.

Mikel, M.A., B.W. Diers, R.L. Nelson, H.H. Smith. 2010. Genetic diversity and agronomic improvement of North American soybean germplasm. Crop Sci. 50:1219-1229.

Mumm, R.H. 2013. A look at seed product development with genetically modified crops: Examples from maize. J. Agricultural and Food Chemistry 61(35): 8254-8259. DOI: 10.1021/ jf400685y

Mumm, R.H., J.W. Dudley. 1994. A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. Crop Sci. 34:842-851.

Mumm, R.H., L.J. Hubert, J.W. Dudley. 1994. A classification of 148 U.S. maize inbreds: II. Validation of cluster analysis based on RFLPs. Crop Sci. 34:852-865.

Panter, D.M., F.L. Allen. 1995. Using best linear unbiased predictions to enhance breeding for yield in soybean: I. Choosing parents. Crop Sci. 35: 397-405.

Rife, T.W., J.A. Poland. 2015. Field Book: An open-source application for field data collection on Android.

Romkaew, J., T. Umezaki. 2006. Pod dehiscence in soybean: Assessing methods and varietal difference. Plant Production Science 9(4): 373-382.

Tanksley, S.D., S. Grandillo, T.M. Fulton, D. Zamir, Y. Eshed, V. Petiard, J. Lopez, T. Beck-Bunn. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative L. pimpinellifolium. Theoretical and Applied Genetics. 92(2):213-224.

Tanksley, S.D., J.C. Nelson. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theoretical and Applied Genetics. 92(2):191-203.

Yan, W. and I. Rajcan. 2002. Biplot evaluation of test sites and trait relations of soybean in Ontario. Crop Sc. 42(1): 11-20.

How to cite this chapter: Mumm, R.H. (2023). Chapter 2. The Process of Cultivar Development: Pure Line Variety. In W. P. Suza, & K. R. Lamkey (Eds.), *Cultivar Development*. Iowa State University Digital Press.

Chapter 3: New Line Development and New Line Evaluation: Single-cross Hybrid

Rita H. Mumm

In Chapter 2, the different concepts and processes, and pertinent considerations in the choice of parents, creation of progeny and materials for testing, and evaluation and selection of desirable lines in cultivar development were examined using a pure line variety of a self-pollinated crop, soybean. In this chapter, we will examine the changes to the process and the product pipeline that would be needed to develop improved hybrid cultivars. The example cited is a single-cross maize hybrid characterized by complete crop uniformity (Fig. 1).

Fig. 1 Hybrid corn. Photo courtesy of Bob Nichols, USDA.

Hybrid Vigor = Heterosis

Hybrid cultivars exploit the advantage of **heterosis**, also referred to as hybrid vigor. The performance of the hybrid is superior to either parent of the hybrid (Fig. 2).

Fig. 2 Parent A (left), Parent B (right), and hybrid crosses AxB and BxA (center). Photos courtesy of A. Melchinger, University of Hohenheim, Stuttgard, Germany.

Heterosis Quantified

Heterosis can be quantified for any given cross and it is typically expressed as a percentage:

$$\frac{100(\text{hybrid - MP})}{\text{MP}}$$

where **MP** is the mid-parent mean (i.e. mean of the parents' performance).

For example: if MP = (5.0 + 4.4) / 2 = 4.7 t/ha and hybrid yield is 11.5 t/ha, then **mid-parent heterosis** =

$$\frac{100(11.5 - 4.7)}{4.7} = 145\%$$

Heterosis can also be quantified for any given cross in terms of performance of the better parent:

$$\frac{100(\text{hybrid - BP})}{\text{BP}}$$

where BP is the mean of the better parent (i.e. max [P1, P2]).

For the case above, **better parent** (also called "high parent") **heterosis** =

$$\frac{100(11.5 - 5.0)}{5.0} = 130\%$$

Heterosis is the Opposite of Inbreeding Depression

The term "heterosis" was coined by George Harrison Shull, who laid the foundation for a more comprehensive understanding of the genetic basis for heterosis based on his work in corn.

He observed that:

- An open-pollinated variety of corn was a complex mixture of hybrids, each plant of a different genotype.
- Inbreeding reduced hybrid vigor; crossing restored it.
- Hybrid vigor is not permanent; hybridization must be done each generation to get the added benefit.
- Inbreds differ in the level of hybrid vigor produced through crossing (combining ability).
- Inbred progeny show more defects than their parents.

George Shull interpreted heterosis as the opposite of inbreeding depression (Shull, 1909). He proposed that the variability among strains undergoing inbreeding, including loss of vigor, was a consequence of segregation and eventual homozygosis of desirable and deleterious alleles. With inbreeding, deleterious alleles are unmasked. And he concluded that heterosis is the product of parent lines compensating for each other. Shull (1908, 1952) applied the concept of heterosis to plant breeding: *'increased vigor, size, fruitfulness, speed of development, resistance to disease and insect pests or to climatic rigors manifested by crossbred organisms as compared with corresponding inbreds as the specific result of unlikeness in the constitutions of uniting parental gametes."*

Dominance Driving Heterosis

Eugene Davenport (1908) proposed the Dominance Hypothesis of heterosis. That is, dominant alleles improve fitness of the hybrid by masking deleterious recessive alleles. Edward M. East (1908) and George H. Shull (1908) independently proposed the Overdominance Hypothesis, which declares that certain combinations of alleles represent an advantage in the heterozygote due to over-expression of the gene involved.

Epistatic Gene Action Implicated

Sewall Wright (1922) established that the relationship between mean performance and the decrease in heterozygosity should be linear, regardless of the degree of dominance (i.e., partial, complete, over-dominant) unless linkage or epistasis is involved.

Several studies demonstrated epistasis as a factor in heterosis but accounted for a relatively small portion of the total variability (approximately 10% in Stuber and Moll (1971)).

Dominant gene action is considered the primary factor underpinning heterosis, although more detail on the mechanism is lacking. Additivity is also implicated.

Genetic Diversity Required

The oldest and most famous heterotic pattern is U.S. Yellow Dent corn.

In 1922, Frederick D. Richey observed that hybrids between varieties of different endosperm types in maize resulted in higher yield than varieties of the same endosperm type, suggesting that genetically or geographically distant parents exhibited increased heterosis in hybrid combination. This led to the development of the Reid Yellow Dent x Lancaster Sure Crop pattern, which was the basis for the predominant heterotic pattern developed in U.S. maize: Stiff Staff Synthetic (SSS) x non-Stiff Stalk (NSS).

We can learn a great deal from this important development.

What is a Heterotic Pattern?

A **heterotic pattern** is a specific pair of heterotic groups whose crosses express high heterosis (and hybrid performance) (Fig. 3). Genetic diversity is necessary but not sufficient.

A **heterotic group** is a group of related or unrelated individuals from the same or different populations, displaying similar combining abilities when crossed with genotypes from other germplasm groups.

Within a heterotic group, there may be a group(s) of individuals that are more genetically similar to one another than to members of the heterotic group at large. Such a group is referred to as **heterotic subgroup**.

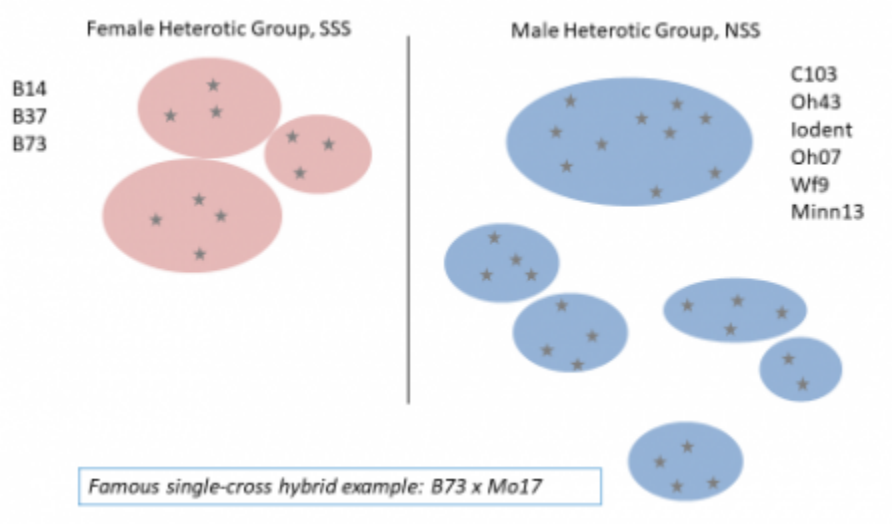


Fig. 3 Heterotic pattern in U.S. maize: Stiff Stalk Synthetic (SSS) x Non-Stiff Stalk (NSS).

Properties of a Heterotic Pattern

A heterotic pattern has the following characteristics:

    - **Genetic diversity** between the two heterotic groups. Diversity is necessary but not sufficient for expression of heterosis. That is, two genetically diverse lines do not necessarily display heterosis in hybrid combination, but two lines that display heterosis in hybrid combination are always genetically diverse.
    - **General combining ability** is expressed between inbreds from opposite heterotic groups.

- **Specific combining ability** is expressed between specific pairs of inbreds from different heterotic groups… and exploited by plant breeders!

Preserving Genetic Diversity Between Heterotic Groups

The heterotic pattern must be taken into account when choosing parents to create breeding populations. The principles of choosing prospective parents with the best performance for key traits and choosing pairs that represent genetically diverse sources of favorable alleles still hold. However, with hybrid cultivars, it is critical to preserve heterosis between heterotic groups. Therefore, breeding crosses are made *within* a heterotic group, not between heterotic groups.

Improving a Specific Single-cross Hybrid

Consider a specific single-cross corn hybrid we will call H_{Now} with parents I_1 (from heterotic group 1) and I_2 (from heterotic group 2).

$$\begin{array}{c} \mathrm{I}_1 \times \mathrm{I}_2 \\ \downarrow \\ \mathrm{H}_{Now} \end{array}$$

Suppose we want to improve this hybrid. How would we go about it?

Improving One Parent of a Specific Hybrid

One of the parents of the hybrid, say I_1, could be crossed to a line from the same heterotic group that represents a source of new favorable alleles. Progeny of this breeding cross could be evaluated for their performance in hybrid combination with a line from the opposite heterotic group (called a **tester**) to make selections (Fig. 4). In this case, a superior progeny advanced to homozygosity would represent a new inbred (I_{New}) which when combined with I_2 results in an improved hybrid, I_{New}, which outperforms H_{Now}.

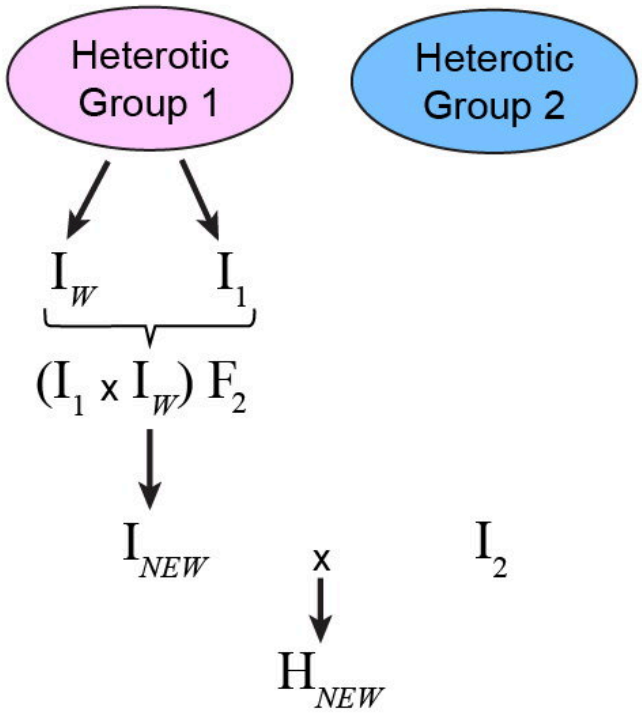


Fig. 4 Process flowchart for improving one parent before creating new progeny for evaluation.

Improving Both Parents of a Specific Hybrid

In practice, breeding crosses are typically made within each heterotic group to improve both sides of the pedigree (Fig. 5). Progeny from each cross are evaluated with respect to performance in hybrid combination with a tester(s) from the opposite heterotic group, especially for traits involving non-additive gene action. The testers can be the parents of H_{Now} or other prominent lines from the opposite heterotic group.

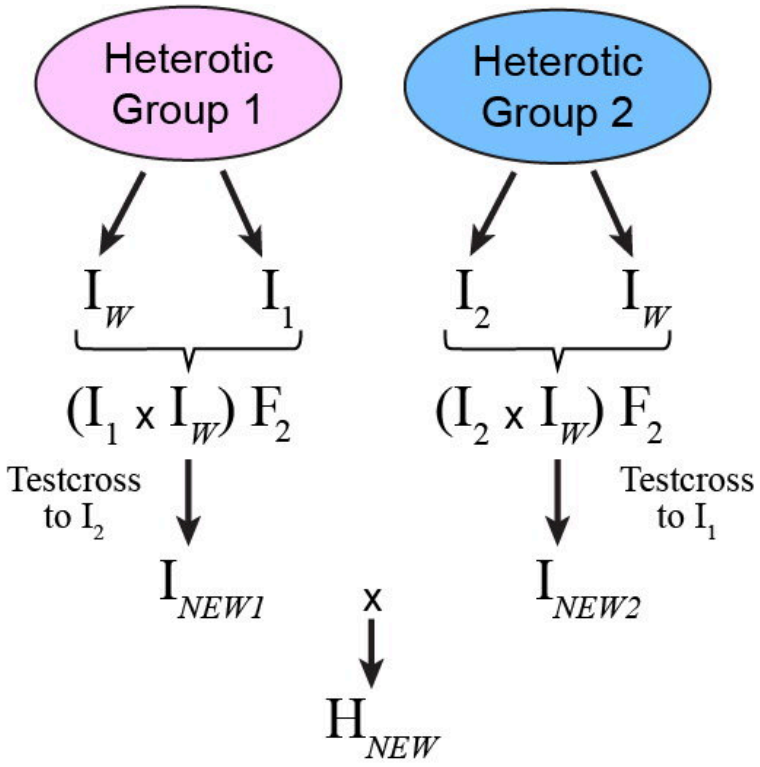


Fig. 5 Process flowchart for improving both parents before creating new progeny for evaluation.

Using a Prospective Source of New Favorable Alleles

Consider a source of favorable alleles for key traits, say I_W, to improve H_{Now}:

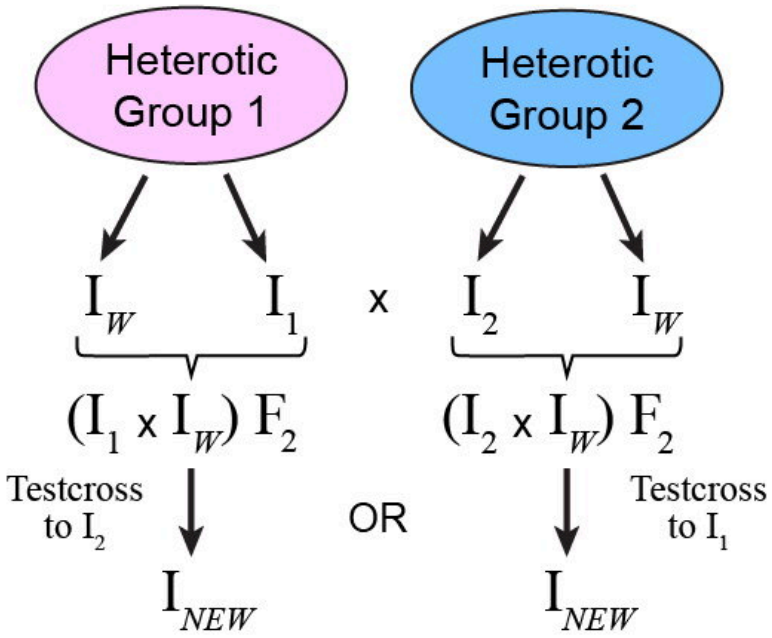


Fig. 6 Flowchart for improving both parents before creating new progeny for evaluation by making double-crosses or testcrossing F_2 progeny.

I_W could be crossed to I_1 to create F_2 families that can be evaluated for their performance in hybrid combination with I_2. A new superior inbred resulting from I_1 x (that is, I_{New}) could be used in combination with I_2 to produce an improved hybrid, H_{New} (Fig. 6).

Alternatively, I_W could be crossed to I_2 to create F2 families that can be evaluated for their performance in hybrid combination with I_1 with the recovery of a new superior inbred.

To which inbred (I_1 or I_2) should I_W be crossed to create a breeding population that will ultimately contribute to development of a new improved hybrid, H_{New}?

This answer is easy if you know the pedigree of inbreds I_W and I_1 and I_2. You would cross I_W to the inbred to which it is more related (that is, more genetically similar) to create a breeding population, thus preserving heterosis between I_{New} and the other parent line.

Assigning New Germplasm to a Heterotic Group

Assuming a particular heterotic pattern has been recognized and utilized, new potential breeding materials can be assigned to a relevant heterotic group through:

- Pedigree analysis
- Geographic inference
- Genetic similarity based on molecular marker profile using cluster analysis (e.g. Mumm and Dudley, 1994)

- Measurement of heterosis
- Combining ability analysis

The latter two are typically conducted using materials created through mating designs such as diallel or partial diallel, or NC Design II.

Establishing a Heterotic Pattern

Separate heterotic groups are essential for the development of hybrid cultivars. If heterotic groups have not yet been defined, there are approaches that can provide guidance on a starting point. A diallel mating design can be used to measure heterosis between pairs of lines (as demonstrated in Eberhardt, 1971). Heterotic grouping can be arbitrarily assigned to maximize heterosis between heterotic groups.

In addition, the phenotypic data can be supplemented by cluster analysis of a larger set of lines based on molecular marker profile (see Chapter 2 for guidance on cluster analysis based on molecular marker profile). The cluster analysis can be used to identify genetically similar materials, with the assumption that genetically similar lines will perform similarly in hybrid combinations. The lines that showed heterosis in hybrid combination can serve to anchor the groupings (clusters) and indicate heterotic group assignments as well as heterotic subgroup membership.

The newly established heterotic pattern can be reinforced and strengthened through the use of breeding methods such as reciprocal recurrent selection. Furthermore, one of the heterotic groups in a newly established heterotic pattern will be chosen to serve as the pool of female hybrid parents and the other as the male. Thus, selection within a heterotic group will also focus on reproductive features important to its role in hybridization. That is, lines in the female pool will likely be improved for seed set whereas lines in the male pool will be improved for traits associated with pollen shed.

Parents of Breeding Crosses Chosen Within a Heterotic Group

In developing improved hybrid cultivars, New Line Development will take place within the context of each heterotic group. That is, parents for breeding populations are chosen within, not between, heterotic groups. To maximize genetic diversity within the breeding population, parents may be chosen from different heterotic subgroups within the heterotic group. For example, the choice of parents for a breeding population to improve the female side of the pedigree in U.S. maize may involve a line from the B14 heterotic subgroup and a line from the B73 subgroup (Fig. 3).

Prospective parents within a heterotic group can be evaluated on the basis of estimated breeding value (EBV) in hybrid combination with members of the opposite heterotic group for traits involving heterosis. Whereas EBVs for pure line varieties consider the merit of lines *per se*, EBVs for prospective parents of improved inbreds for hybrid cultivars take complementarity into account. (For a comparison of examples, see Bernardo, 2010.)

Classes of Loci

For any given prospective parent, $\mathbf{I}_W$, the value of its contribution to improving a given hybrid, $\mathbf{I}_1$ x $\mathbf{I}_2$, can be assessed based on the methodology proposed by Dudley (1987).

To increase the frequency of favorable alleles in $\mathbf{H}_{Now}$, certain classes of loci are key. With inbred parents, there are eight classes of loci relative to $\mathbf{I}_1$, $\mathbf{I}_2$, and $\mathbf{I}_W$, as shown in Table 1.

Table 1 Classes of loci proposed by Dudley, 1987.

Class of loci	$\mathbf{I}_1$	$\mathbf{I}_2$	$\mathbf{I}_W$
A	++	++	++
B	++	++	– –
C	++	– –	++
D	++	– –	– –
E	– –	++	++
F	– –	++	– –
G	– –	– –	– –
H	– –	– –	– –

Because all lines are homozygous, only one allele is present at each locus; "**+**" is used to indicate the dominant (favorable) allele, and "**–**" to indicate the recessive allele in Tables 1, 2, 3, and 5.

Which Classes of Loci are Important?

Which classes of loci are important to developing an improved hybrid?

With Classes A and H, all lines are invariant; no new favorable alleles are introduced.

If dominance is complete (a=1), then no need for improvement at Class C or D if $\mathbf{I}_2$ is to be improved, or at Class E or F if $\mathbf{I}_1$ is to be improved.

With Class G, $\mathbf{I}_W$ has new favorable alleles not present in either $\mathbf{I}_1$ or $\mathbf{I}_2$, which is imperative.

Potential Gain / Potential Loss

If $\mathbf{I}_W$ is crossed to $\mathbf{I}_1$, there is potential for loss with Class D loci, and the potential for gain with Class G loci in $\mathbf{H}_{New}$ (Table 2).

Table 2 Status of classes of loci with improvement of I_1.

Class of Loci	I_1 x I_W	I_2
A	++	++
B	++	++
C	++	—
D	+-	—
E	—	++
F	—	++
G	+-	—
H	—	—

If I_W is crossed to I_2, there is potential for loss with Class F loci and the potential for gain with Class G loci in H_{New} (Table 3).

Table 3 Status of classes of loci with improvement of I_2.

Class of Loci	I_1	I_2 x I_W
A	++	++
B	++	++
C	++	– –
D	++	– –
E	– –	++
F	– –	+–
G	– –	+–
H	– –	– –

Clearly, I_W needs to be a source of new favorable alleles (Class G), but the potential loss through Class D or F must be considered as well.

Dudley (1987) devised a methodology to evaluate potential new parental sources to determine which have the highest probability of producing an improved hybrid H_{New} when used to create breeding populations with either I_1 or I_2. The method also provides guidance on which parent of H_{Now} to improve and whether to create F_2 or BC_1 families as a starting point.

General Principles of Choosing Parents

In general, the principles of choosing parents do not change in working with a hybrid cultivar (versus a pure line variety):

- Need sources with a *high frequency of favorable alleles* for the traits defined in the product target.
- Need a *diversity of favorable alleles* between the parents of the breeding crosses so there is genetic variability among the progeny comprising the breeding population.

In developing hybrid cultivars, there is also a need to preserve heterosis between the heterotic groups. Therefore, breeding crosses are made *within* heterotic groups and progeny are evaluated in terms of **testcross** performance to meet the need to build *complementarity between heterotic groups*.

Ultimately, improved hybrids illustrate strong specific combining ability.

Controlled Pollination

Of course, breeding crosses are made in such a way as to strictly control the parental contributions of **pollen** and **ovule**.

One major difference in New Line Development and New Line Evaluation between hybrid cultivars and pure line varieties is the need for controlled pollinations once the breeding populations have been formed. As we saw with soybean, once the breeding cross is made, all other pollinations are self-pollinations. With soybean, self-pollination takes place even before the **complete flower** opens.

With the development of hybrid cultivars, there are situations where seed produced through self-pollination is not the objective. Thus, a means to perform controlled pollinations with a specific pollen parent is needed on a fairly large scale.

Consider steps for controlled pollination in maize… see the following YouTube video by Tracy and Kaeppler at the University of Wisconsin for a demonstration:

Breeding Strategies

Let's consider breeding strategy since this is, in large part, a function of choice of the base germplasm used and choice of parents for breeding crosses.

The *pedigree breeding method* is popular with hybrid cultivar development (see Chapter 2 for detail on the pedigree breeding method). Elite lines are crossed to develop a breeding population. However, in developing hybrid cultivars, progeny will be evaluated for testcross performance for any traits involving heterosis.

If the base germplasm is not particularly elite for yield performance and other non-additive key traits, *recurrent selection* methods may be practiced as a means to increase the frequency of favorable alleles within a heterotic group.

Recurrent Selection

Recurrent selection involves the systemic development and improvement of a breeding population.

Individuals in the population are evaluated with respect to a trait(s) of interest and selected lines are intermated to form the next cycle.

In practicing recurrent selection, there are a number of mating designs that can be used to create materials for testing. Selection can take place on an individual plant basis or on the basis of family structure (i.e. half-sib families, full-sib families, S_1 families), depending on the heritability of the trait(s) for which the population is to be improved.

The improved population can be used as a source of new inbreds. With any cycle, top-performing lines can be spun out for use as parents of improved hybrid cultivars and/or used as parents to develop breeding populations (e.g. implementing pedigree breeding).

Iowa Stiff Stalk Synthetic, a synthetic variety developed in 1933-34 by George F. Sprague during his tenure at Iowa State University (Sprague, 1983), is a famous case of recurrent selection. Iowa Stiff Stalk Synthetic (BSSS) was developed with 16 lines chosen for stalk strength that were intermated to form a population that was random mated for an unknown number of generations. Several important female inbreds were spun out of various cycles: B14 from Cycle 0, B37 from Cycle 0, B73 from Cycle 5, and B84 from Cycle 7. BSSS became the basis for the female side of the pedigree in U.S. commercial maize germplasm (referred to as Stiff Stalk Synthetic, SSS), and its prominence is still noted today.

Reciprocal Recurrent Selection

In the case of hybrid cultivars, **reciprocal recurrent selection** may be used to simultaneously improve two populations (one from each heterotic group), boosting the frequency of favorable alleles in a complementary fashion. Thus, each population is improved with respect to the other.

The following example of reciprocal recurrent selection involves three steps per cycle:

1. In the first season, a plant in Population 1 (P1) is selfed (to produce S_1 seed) and crossed to several random plants in Population 2 (P2). The seeds harvested from the P2 plants are bulked to form a P1 half-sib family. The self and cross procedure is also done in P2.
2. In the second season, the half-sib families in P1 (crossed to P2) and P2 (crossed to P1) are evaluated in performance tests.
3. In the third season, the selfed seeds from the best plants in P1 are grown and the plants are intercrossed to form the next cycle. The selfed seeds from the best plants in P2 are likewise used in recombination separately from P1.

GCA and SCA

With reciprocal recurrent selection methods, selection within each population is based on general combining ability (GCA). Variance associated with GCA measures variation among progeny with one common parent (i.e., half-sib families).

Lines representing newly improved inbreds may also be selected based on specific combining ability (SCA). Variance associated with SCA involves variation associated with female x male line interaction.

Any selection scheme based on testcross performance with the opposite heterotic group and applied to both sides of the pedigree can be considered a form of reciprocal recurrent selection. This includes pedigree breeding with selection of progeny based on testcross performance in combination with a tester line from the opposite heterotic group.

Mating Designs to Measure GCA and SCA

There are a number of mating designs which can be used to measure GCA and SCA effects, the most common of which is the **diallel**.

A **complete diallel** involves every possible paired combination of lines. Reciprocal crosses and parent lines may or may not be included in the phenotypic evaluation (Griffing, 1956). A total of $\frac{p-1}{2}$ crosses are made with *p* lines in a complete diallel where reciprocal crosses are bulked. To minimize the number of crosses needed for evaluation of a large number of lines, parents may be assigned to *sets* which are essentially mini-diallels.

In addition to providing estimates of GCA and SCA, a diallel can be used to compute estimates of heterosis. Thus, diallel can be used in establishing a new heterotic pattern; lines showing the greatest heterosis in hybrid combination are arbitrarily assigned to opposite heterotic groups.

The **NC Design II** is another mating design useful for obtaining estimates of GCA and SCA effects.

Vital Need for a Testing System Based on Hybrid Performance

A testing system based on hybrid performance is vital to the development of improved hybrid cultivars. After all, the improved cultivar will be exhibiting its hybrid performance to farmers and other stakeholders in the value chain.

Sufficient, accurate data on hybrid performance is essential to making a correct decision about product launch.

Choice of Tester

Before examining a testing regime suited to New Line Development and New Line Evaluation for hybrid cultivars, let's first consider the choice of a tester to produce testcross seed for evaluation…

Since testcross evaluation is used to assess combining ability (and to estimate breeding values), the challenge is to find a tester that provides discrimination among progeny in a breeding population in keeping with the purposes of selection.

Matzinger (1953) defined a desirable tester as one that combines the greatest simplicity of use with the maximum information on the performance to be expected from tested lines when used in other hybrid combinations or grown in other environments.

Requirements of a Tester

An ideal tester maximizes differences among the genotypes being tested. That is, an ideal tester maximizes the variance among testcrosses:

$$V_{TC} = \frac{1}{2}(1+F)pq\left[a + d(q_T - p_T)\right]^2$$

The equation indicates that V_{TC} is a function of:

- Allele frequencies in the population, ***p*** and ***q.***
- Allele frequencies in the tester, $\boldsymbol{p_T}$ and $\boldsymbol{q_T}$.
- Level of inbreeding in the population, ***F.***
- Levels of dominance, ***d.***
- The value, *a,* which is equal to half the difference between genotypic values of the two homozygotes (i.e. additive effect) across loci affecting trait performance.

Note that for any level of dominance, the quantity

$$\left[a + d(q_T - p_T)\right]^2$$

is maximized when $\boldsymbol{q_T} = 1$. That is, the tester is fixed for the recessive allele at most underlying loci.

The recessive allele from the tester does not mask the effect of a favorable, dominant allele contributed by the progeny. However, it also suggests poor performance *per se* of the tester line.

Also note that if ***d*** =0, then the tester has no effect. However, most loci involved with the expression of heterosis are expected to have some level of dominant gene action.

Comparison of Tester Types

Hallauer and Lopez-Perez (1979) compared five types of testers in maize for their value in assessing lines from Iowa Stiff Stalk Synthetic (BSSS) at early (S_1) and late (S_8) stages of inbreeding. The testers were compared as to their testcross variance, V_{TC}, and mean performance, $\boldsymbol{\mu_T}$, for grain yield, a highly heterotic trait (Table 4).

Testers evaluated in combination with BSSS included:

- The BSSS population itself.
- BC13(S)C1, an improved version of the BSSS population.
- BSSS-222, a low-yielding inbred selfed out of BSSS.

- B73, a high-yielding inbred selfed out of BSSS Cycle 5.
- Mo17, a high-yielding inbred line from the opposite heterotic group.

Table 4 Comparison of five types of testers. Data from Hallauer and Lopez, 1979.

Tester	Description	S_1 V_{TC} (t/ha)	S_8 V_{TC} (t/ha)	S_1 μ_T (t/ha)	S_8 μ_T (t/ha)
BSSS	Population itself	0.18	0.42	5.79	5.69
BC13(S)C1	Improved BSSS population	0.11	0.34	6.95	6.81
BSSS-222	Poor BSSS Inbred	0.22	0.39	6.03	5.89
B73	Elite BSSS inbred	0.04	0.26	7.29	7.21
Mo17	Elite non-BSSS inbred	0.26	0.30	7.81	7.78

Conclusions:

- More differentiation among progeny was possible with S_8 vs S_1 testcrosses (i.e., greater *F* in the progeny highlighted the allele frequencies) despite consistency in performance levels.
- The poor performing testers generally led to higher V_{TC} (e.g., BSSS-222 vs B73; BSSS vs BC13(S)C1). Note that all of these testers are related to BSSS.
- In early testing, V_{TC} was maximized by the Mo17 tester, the elite line from the opposite heterotic group.

Clearly, the Mo17 tester is highlighting differences among BSSS lines at loci for which Mo17 is deplete of favorable alleles! Also, the Mo17 tester provided the best mean performance for yield in both early and late testing. This is a tremendous advantage because a poor tester cannot be considered as a potential parent of an improved hybrid; an elite line from the opposite heterotic group can.

Poor vs. Elite Tester

Thus, there are two types of testers that can maximize V_{TC}:

1. A poor-performing tester, presumably due to a low frequency of favorable alleles
2. An elite inbred tester from the opposite heterotic group, especially lines that could potentially serve as parents of an improved hybrid.

With the second option, yield trials to evaluate testcross performance are identifying the best lines for one side of the pedigree, and the yield trials also capitalize on SCA to serve in identifying top-performing hybrid combinations as potential new cultivars!

At the Genotypic Level

Let's dive deeper...

Consider the following classes of loci in S_8 progeny resulting from a breeding cross (Dudley, 1984) (Table 5).

Table 5 Classes of loci in S8 progeny in relation to poor and elite testers. (Dudley, 1984).

Class of Loci	Breeding Inbred 1	Cross Inbred 2	Poor Tester	Elite Tester
i	++	++	—	++ or —
j	++	—	—	—
k	—	++	—	—
l	—	—	—	++

The poor tester has recessive alleles at most or all classes of loci, whereas the elite tester from the opposite heterotic group is homozygous for the dominant favorable allele at Class '*l* ' loci for which the progeny are depleted of favorable alleles. The elite tester may be homozygous for the dominant favorable allele at Class '*i* ' loci, but these alleles don't contribute to better performance since the progeny are already fixed for the favorable allele.

Progeny are segregating for Classes '*j* ' and '*k*' only. Therefore, these are the loci affecting V_T . Since neither tester is contributing to these loci, V_{TC} is the same with both testers.

However, the elite tester from the complementary heterotic group contributes at Class 'l' loci, increasing μ_T.

Take-Home Message

Bottom line: an elite inbred(s) from the complementary heterotic group is an ideal tester.

All progeny are compared on the same basis using a common tester or a common set of multiple testers. It will highlight real genetic differences among progeny under evaluation and, thus, serve to optimize the response to selection!

In addition, potential new single-cross hybrid cultivars may be detected in the evaluation process.

A Word About Producing Testcross Seed

To produce testcross materials for trials, progeny can be crossed to a tester from the opposite heterotic group through the use of a **topcross** mating design. Generally, a topcross design is carried out in an isolated field; families of the breeding population are grown in separate rows with the male flower removed before pollen shed. The tester, grown in rows placed intermittently through the field, serves as the pollen parent in the testcross, with pollen transmitted to progeny rows through wind, insects, or other means. The field isolation prevents stray pollen from contaminating the testcross seed.

Fig. 7 Hand pollination to make testcross seed. Photo courtesy of Iowa State University.

Another alternative is to produce testcross seed through **paired-row crosses**. Each progeny (or progeny family) is represented in one row, with the tester grown in the adjacent row. Silks are protected to prevent inadvertent pollen ahead of hand-pollination and pollen is collected and distributed manually (Fig. 7); thus, field isolation is not required. Seed produced on plants in either of the paired rows can be bulked to represent the testcross seed associated with that progeny.

Product Target Example: Maize Single-cross Hybrid

Consider the following product target for a breeding program focused on single-cross maize:

Product Target:
Develop a medium maturity, white corn hybrid for dryland production in western South Africa that outperforms Hybrid X for grain yield and has good standability, a strong disease resistance package, and features Bt and R traits.

The **product target** describes the “what” and the “for where.”

Target market region

Where is the improved hybrid intended for? Describe the target market region specified in this product target.

Click here to reveal the answer
Dryland production regions of western South Africa.

Product target environment

The target market region also implies what population of environments?

Click here to reveal the answer
Environments representing geographic location and also the season, maturity zone, altitude, soil types and farmer practices.

What characteristics does this product target imply?

- White grain.
- High grain yield.
- Tolerance to moisture stress levels common to non-irrigated production.
- Good **standability** (minimal/no **stalk lodging** or **root lodging**).
- A strong "disease package" for this region might include resistance to.
  - Grey leaf spot (*Cercospora zeae-maydis* and *Cercospora zeina*)
  - Common rust (*Puccinia sorghi*)
  - Diplodia ear rot (*Stenocarpella maydis*)
- Medium **maturity**.
- **Bt event** (e.g., transgenic event which imparts resistance to stalk borers).
- **RR event** (e.g., transgenic event which imparts tolerance to glyphosate herbicide such as Roundup®).

Possible Secondary Traits to Collect

What characteristic(s) would be a good indicator of tolerance to moisture stress in maize?

Yield evaluations will be conducted in the target market region, which will provide an indication of response to any level of drought encountered in that environment.

In addition, another indicator could be considered: anthesis-silking interval (ASI). ASI is the number of days between the date when 50% of plants are shedding pollen and the date when 50% of plants are showing emerged silks. ASI is an external indicator of the metabolics underlying stress tolerance. Large ASI in maize is a sign of reduced partitioning to the ear, resulting in slow spikelet growth which affects kernel number. ASI is highly correlated with **barrenness** which, like kernel number, is a contributing factor to grain yield. ASI is a clear indicator of stress at flowering, which is the most vulnerable time for yield loss due to drought under "occasional drought." ASI not only signals moisture stress, it is a good indicator of other abiotic stresses such as nitrogen deficiency.

Use of ASI

Is the use of ASI to indicate tolerance to moisture stress an example of a secondary trait?

Click here to reveal the answer
Yes!

Measuring Performance of Desired Characteristics

The specific levels of the desired characteristics, the selection thresholds, and the way the characteristics will be measured can be detailed after the target characteristics. Any protocols to be used in trait evaluation can be specified.

For example (Table 6):

Table 6 Characteristics, measurement standards, and selection thresholds associated with the stated product target for maize hybrid.

Characteristic	Measurement Standard	Threshold Level / Range
White grain	Visual; 1 = white	1
High grain yield	Machine harvest; grain weight on a 12.5% moisture basis, expresses per unit of land	10% greater than Hybrid X
Tolerance to moisture stress	ASI (anthesis-silking interval) calculated as the difference between date of 50% silk emergence. Compute the number of days from planting to anthesis and number of days from planting to silking, from record of anthesis date and silking date.	$-2 \leq x \leq 2$
Stalk lodging	Estimated % affected plants per plot	≤ 5%
Root lodging	Estimated % affected plants per plot	≤ 5%
Gray leaf spot	1-9 disease scoring scale, 1=no infection, 2=very low infection, 3=low infection. Protocol for pathogen inoculation and disease screening.	Score ≤ 2
Common rust	1-9 disease scoring scale, 1=no infection, 2=very low infection, 3=low infection. Protocol for pathogen inoculation and disease screening.	Score ≤ 3
Diplodia ear rot (*Stenocarpella maydis*)	1-9 disease scoring scale, 1=no infection, 2=very low infection, 3=low infection. Protocol for pathogen inoculation and disease screening.	Score ≤ 2
Medium maturity	Growing degree days from planting to physiological maturity	110-115 day range (warm areas)
Bt event (YieldGard II)	MON89034 expression	Pre-determined level of Lepidopteran resistance
RR2 event (Roundup Ready 2)	MON88017 expression	Pre-determined level of glyphosate tolerance

Note that Bt and RR2 events require licensing by the breeder's organization. The events listed here are examples.

Example Commercial Maize Hybrid Improvement Program

Consider a suitable testing regime to identify hybrids that meet the example product target, starting with a commercial maize hybrid breeding program as an example. Note that this program serves a market region with one growing season (Summer) (Table 7):

Table 7 New Line Development and New Line Evaluation in a commercial breeding program for hybrid maize.

Season[1]	Activity
Summer 0	Make breeding crosses
Winter 0	Self or BC
Summer 1	1. Grow 40 F_2 or BC_1 populations (i.e. S_0 generation) with 200+ plants each 2. Select S_0 plants in each population based on plant type, disease and insect resistance, marker genotype, etc. 3. Self and testcross 100 selected S_0 plants in each population to an inbred tester
Summer 2	1. Discard ≤5 F_2 or BC_1 populations based on performance data on the parents, obtained from the previous year 2. Grow 3500 S_0 testcrosses in unreplicated trials at 2-4 locations 3. In each S_1 family, self ≥3 plants to obtain two S_2 subfamilies per S_1 family. Discard S_1 families that appear inferior. 4. Select the best 350 families based on their S_0 testcross performance.
Winter 2	1. Cross 700 S_2 families to each of two inbred testers 2. Self S_2 families to obtain S_3 families
Summer 3	1. Evaluate 1400 S_2 testcrosses in unreplicated trials at 6-10 locations 2. Select 8-15 S_3 families based on their S_2 testcross performance; code lines
Winter 3	Self the selected S_3 families to obtain S_4 seeds of new inbreds
Summer 4	1. Cross each new inbred with 6-10 elite inbreds 2. Self the S_4 lines to obtain S_5 seeds of new inbreds
Summer 5	1. Yield trials of experimental hybrids at 15-40 locations 2. Self the S_5 lines to obtain S_6 seeds of new inbreds
Summer 6	1. Yield trials of advanced hybrids at 20-75 locations 2. On-farm strip tests (i.e. 150-300 m^2 plots) at 30-500 locations
Summer 7	On-farm strip tests of pre-commercial hybrids at 50-1500 locations
Fall	Release 0-2 new hybrids

[1] Number after season indicates the year in the development pipeline. Note: This program serves a market region with one growing season per year (Summer). Winter indicates off-season activities.

Salient Features

We observe the following with respect to the Example Commercial Maize Hybrid Improvement Program:

- The main focus of testing is grain yield in this example; high grain yield is obviously the trait of the utmost priority in producing new hybrids (in this case).
- Selection for highly heritable traits is performed in early generations (selection among individual S0 plants is indicated in this example).
- A large number of breeding populations are created, with about 50-100 plants per population. This would include breeding crosses made with bot heterotic groups (e.g., 40 female populations and 40 male populations).
- All yield trials are conducted using testcross materials.
- Yield trials are initiated with a large number of S_0 testcross lines, each representing a different S_0 family.
- Selection intensities are high (only a small proportion of the tested families are advanced to the next generation of testing).
- The number of elite inbred testers crossed to each family under evaluation for yield increases as the number of tested families decreases.
- The number of locations for yield testing increases as the number of tested families decreases.
- Lines that are selected after Summer 3 yield trials (second year of yield testing) are **coded** as "experimental" and progressed as S_4's to wide-area testing the following growing season; once coded lines are coupled with elite inbred testers, it is the hybrid that is under selection, not the S_4 line.
- Wide-area testing includes "on-farm trials" managed by potential seed customers in the second year.
- Overall, there are at least four years of comprehensive performance testing: four seasons of "research" yield trials, two of which are wide-area testing.
- A pre-commercial stage to acquaint farmers with potential new products is featured in the fifth year of testing; this year could also include National Performance Trials required for registration of a new variety.
- Because maize plants can be self-pollinated and the tassel is used for cross-pollination at the same time, nurseries require a setup that facilitates producing both types of seed.

Adaptation Needed to Meet the Stated Product Target

Given this example of New Line Development and New Line Evaluation from commercial maize, what modifications in the breeding and testing regime are needed to fit the evaluation and selection required to achieve the example product target?

Note that after seven years of hard work and significant resource expenditure, it is possible to wind up with zero new, improved hybrids! *How will your pipeline design of New Line Development and New Line Evaluation minimize the potential for failure?*

Male Sterility Required for Hybrid Seed Production

An important consideration in the development of hybrid cultivars is the need to affect male sterility during hybrid seed production in a cost-effective manner. Controlled pollination on a large scale is essential to ensure that the hybrid seed being produced for the farmer is a product of the intended cross between the elite **female inbred** (from the female heterotic group) and elite **male inbred** (from the male heterotic group). In particular, prevention of self-pollination by the female inbred plants maximizes the probability that the pollen parent is the desired male inbred in a wind-pollinated crop such as corn.

Fig. 8 Typical arrangement of female and male rows in hybrid production of corn single-cross cultivar. Photo from Hollowaydetasseling.com.

Hybrid production fields are arranged to include a predetermined ratio of female inbred (i.e. seed parent) rows to male inbred (i.e., pollen parent) rows (Fig. 8). Common ratios are 4:1, 4:2, 6:2, depending on the pollen dissemination range. Hybrid production fields are grown in isolation to avoid the intrusion of foreign pollen.

Ways to Control Pollination

Control of pollination can be facilitated physically through:

- Separate female/male plants (i.e., the crop species is **dioecious**)
- Manual or mechanical means of removing male flowers or male reproductive structures. In corn hybrid seed production, physical removal of the male flower of the maize plant, referred to as “detasseling”, requires a large short-term labor force (Fig. 9).

Fig. 9 Hand detasseling as a quality check to mechanical detasseling. Photo from www.kokomotribune.com.

Alternatively, male sterility (MS) systems in the plant can be utilized to prevent pollen shed in female inbred parent, which can lead to selfed rather than hybrid seed.

There are several types:

- Genic MS.
- Environmentally-induced MS.
- Cytoplasmic MS.
- Transgenic MS.
- Chemically-induced MS.

Genic Male Sterility

Genic MS is controlled by nuclear genes. It is often determined by a recessive allele at a single locus.

Genic MS was first reported by L.A. Eyster (1921) in maize. Today nearly all identified nuclear MS genes in maize have been mapped. Furthermore, MS can be affected by mutation.

Examples of genic MS include:

- The *ms45* allele in maize: this recessive allele in a homozygous state results in MS, whereas fertility is restored by a single copy of the dominant allele, *Ms45*.
- The suppressor-mutator (Spm) transposable element system (Brar et al. 1987).

Environmentally-induced Male Sterility

Some nuclear MS genes can be induced under certain environmental conditions. Thermo-sensitive and photoperiod-sensitive nuclear genes have been used in rice, for example, to facilitate hybrid production. This enables the MS female inbred lines to be maintained through self-pollination in environments not conducive to MS induction, while hybrid production is carried out under conditions that promote MS.

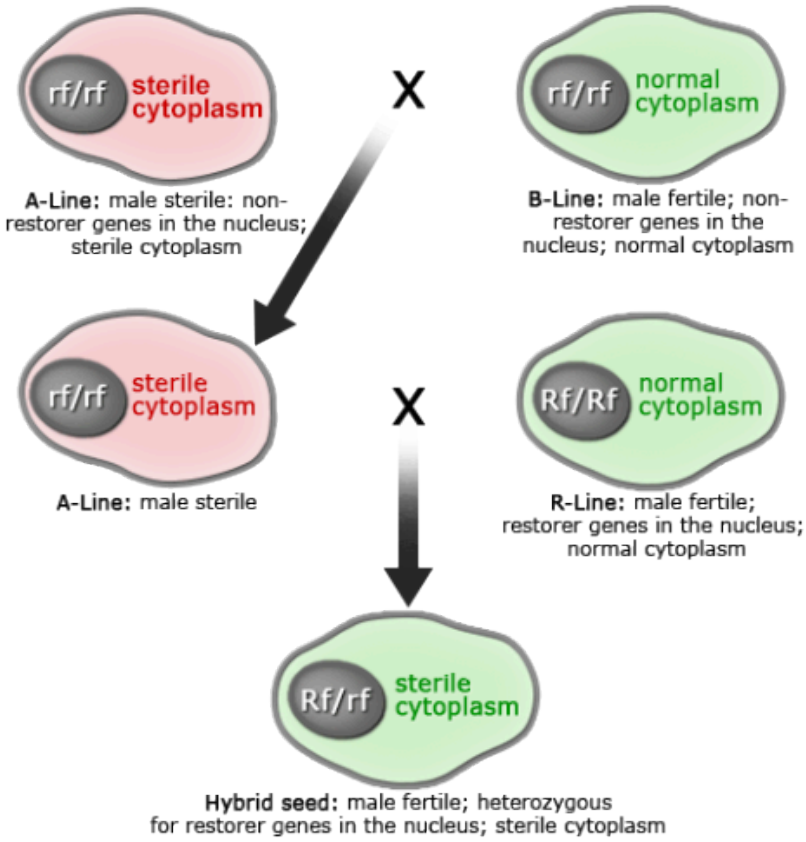
This type of **two-line system** (male-sterile female line, restorer line) is not without issues as weather conditions can be unpredictable and inconsistent. Temperature fluctuations may cause "**leaks**," where some viable pollen is produced, or even complete reversal to a fertile state.

Cytoplasmic Male Sterility

Compared to the normal maize inbred, the CMS-version of inbreds does not exert anthers and therefore does not shed pollen (Fig. 10), a trait that is exploited in hybrid seed production.

Fig. 10 Normal tassel (left) and CMS-version (right) of maize inbred. Note there are no anthers exerted or pollen shed from the CMS version. Photo courtesy of Iowa State University.

Cytoplasmic male sterility (CMS) has been observed in more than 150 species. It is transmitted exclusively by the female parent as a function of the mitochondrial genome. Restoration of fertility relies on nuclear genes that suppress cytoplasmic dysfunction. Each MS cytoplasm has its own “key” to restore fertility in the form of nuclear gene(s) referred to as *Rf* (Restorers of fertility) (Fig. 11).



Fig, 11 Hybrid seed production using the cytoplasmic CMS with Rf gene system. Image courtesy: Iowa State University.

Some examples of CMS systems in maize include:

- **CMS T-type:** Dominant *Rf1* and *Rf2* genes, both required.
- **CMS S-type:** Dominant *Rf3* gene.
- **CMS C-type:** Dominant *Rf4* involved, maybe others.

CMS has been observed in more than 150 plant species and similar CMS systems have been detected in many crops.

Maintaining the CMS Inbred Line

With CMS systems, the need to maintain the MS inbred is accomplished through the use of a **maintainer** line, which is genetically identical to the MS inbred except that it does not have the MS cytoplasm. Both the MS inbred and its maintainer line harbor the recessive MS restorer allele (denoted *r*). When the MS female inbred is crossed to the male inbred containing the nuclear fertility restorer gene (denoted *R*) in hybrid production, the resulting F_1 seed for farmers is male fertile despite having the MS cytoplasm.

Note that this system relies on certain genetic elements on *both* sides of the hybrid pedigree. Such a system is referred to as a **three-line system** (CMS line, maintainer line, restorer line) (Fig. 12).

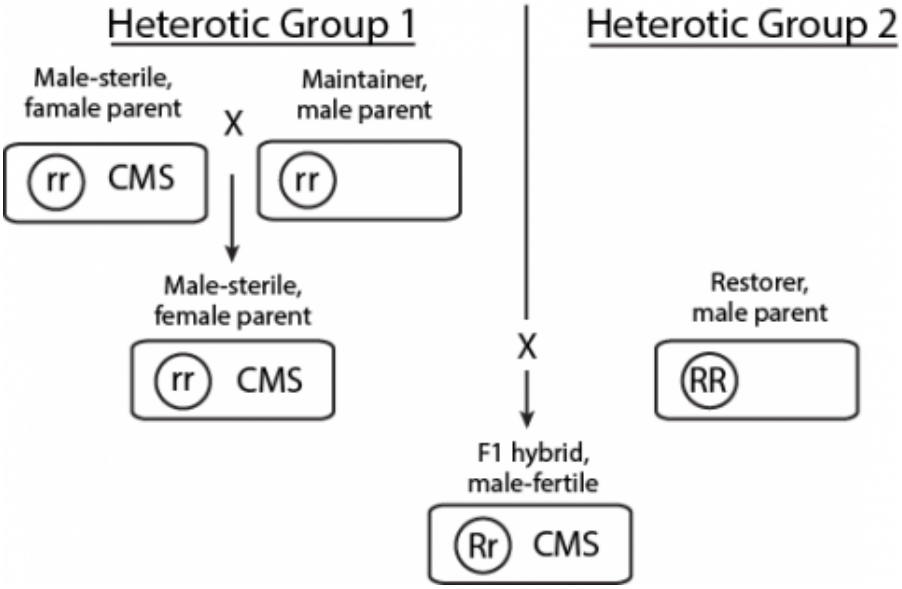


Fig. 12 Schematic for use of R restorer of fertility gene in hybrid production in the CMS system.

Choosing a CMS System

When considering the use of a CMS system, it is critical to compare options and explore the possibility of the negative impact of aberrant mitochondrial genes on hybrid performance.

Fig. 13 Southern corn leaf blight symptoms.
Photo courtesy of David B. Langston, University of Georgia.

The U.S. corn crop was nearly wiped out in 1970 due to the widespread use of T-CMS, which is susceptible to Southern corn leaf blight (Fig. 13) caused by pathogen *Bipolaris maydis (*also known as *Cochliobolus heterostrophus).*

Transgenic Male Sterility

Some systems of male sterility have been created through genetic engineering. For example, DuPont Pioneer (now Corteva Agrisciences) has employed an innovative proprietary Seed Production Technology (SPT) system involving transgenic event DP-32138-1 in maize (Commuri et al, 2009), which relies on the *zm-aa1* gene to render pollen inviable (Fig. 14). The event also includes an *Ms45* gene, which restores fertility in an *ms45/ms45* background as well as a red fluorescent protein (*DsRed2*) gene, which serves as a marker for transgenic seeds. The transgenic male-fertile maintainer line is used to produce non-transgenic seed of the female parent of the hybrid (transgenic seed is also produced due to segregation for the event and this seed is used to propagate the maintainer line). The non-transgenic MS female plants are used in hybrid production to create non-transgenic hybrid seed. Fertility is restored in the F_1 seed through its *Ms45/Ms45* male parent. Thus, the transgenic element of the system does not enter the value chain or the food system.

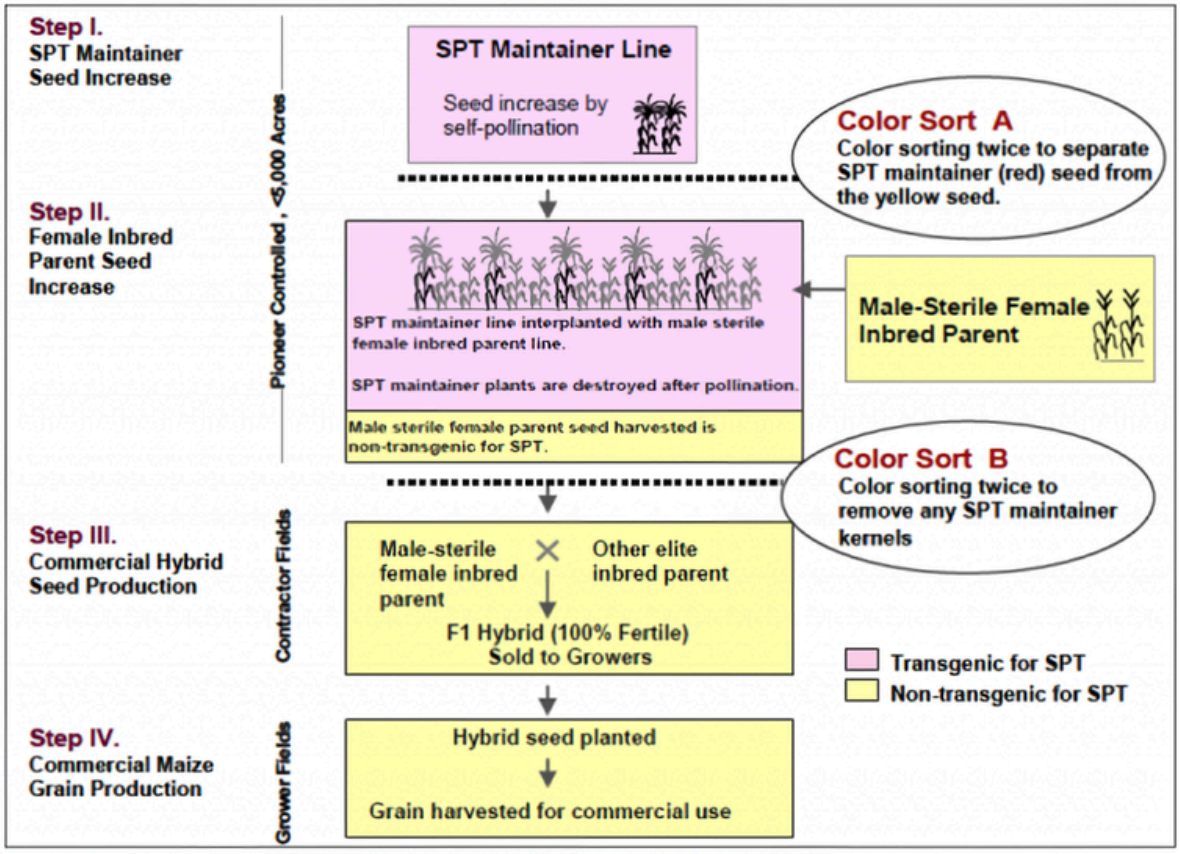


Fig. 14 Steps involved in the deployment of SPT 32138. From USDA APHIS.

A transgenic MS system in rice utilizes a mutant allele of nuclear gene *OsNP1* (*Oryza sativa No Pollen 1*), which encodes a regulator gene controlling tapetum degeneration and pollen exine formation. The *osnp1* mutant exhibits male sterility, which is insensitive to environmental conditions (Chang et al., 2016).

Chemically induced Male Sterility

MS has been induced in a number of crops including rice, sugarbeet, wheat, cotton, rapeseed, canola, cucurbits, tomato, and onion using **chemical hybridization agents** (CHAs) such as male gametocides (e.g., gibberellic acid, maleic hydrazide, naphthalene acetic acid, ethereal). Many CHAs, including classes of sulfonylureas and imidazolinones, inhibit acetohydroxyacid synthase (AHAS), an important enzyme in amino acid biosynthesis. A MS system in rapeseed utilizes a sulfonylurea-resistant mutant to maintain normal male fertility in the pollen parent during hybrid production despite herbicide application to the entire field (Li et al., 2015).

Chemically induced MS can be coupled with transgenic MS systems as a means to "switch on" expression at a desired time. Monsanto's proprietary MS system in maize based on event MON 87427 (Breeze et al, 2010) utilizes glyphosate herbicide (e.g., Roundup®) as a means to induce MS just prior to tassel development. The seed parent in hybrid production has been converted to MON 87427 which protects all plant tissues except tapetum cells and pollen grains against the effect of glyphosate. Thus, herbicide application affects only these tissues. Note that the pollen parent in hybrid production must be protected against damage from glyphosate application through incorporation of a glyphosate-tolerant event which is expressed in all plant tissues.

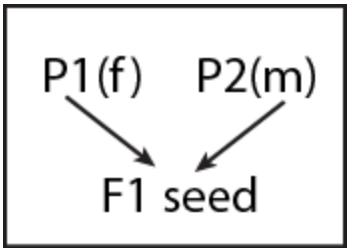
The key with use of chemically induced MS is that female fertility is not affected.

Incorporating MS System Elements Through Breeding

There are many options available to facilitate hybrid production through controlled pollination. MS systems are key to producing genuine F_1 seed of the improved inbreds the breeder has worked diligently to develop.

Although hybrid production to produce F_1 seed for farmers takes place downstream of New Line Development, New Line Evaluation, and Trait Integration, the elements of the MS system used must often be integrated through breeding efforts. Thus, choice of a MS system is a decision affecting the design of the product pipeline. Introgression of MS system elements will be discussed in more depth in Chapter 5 on Trait Introgression.

Single-cross Hybrids



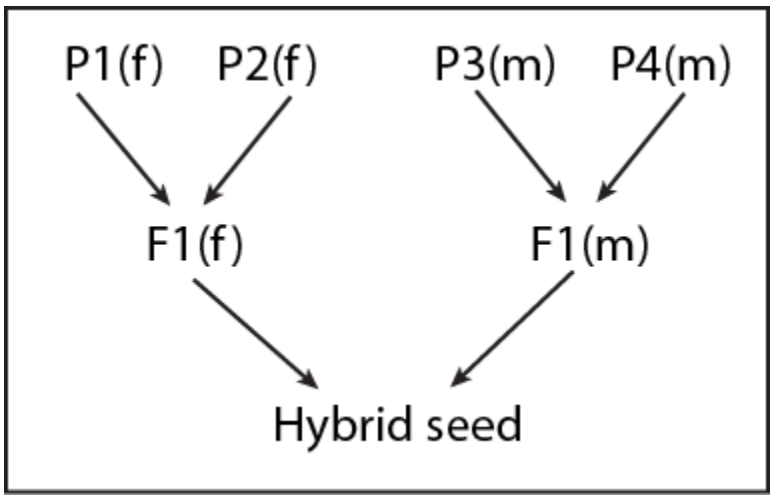
This chapter has focused on the design of New Line Development and New Line Evaluation for improvement

of hybrid cultivars. The focus has been on single-cross hybrids whereby F_1 seed is produced by crossing two genetically unrelated inbred lines, one from the female heterotic group and one from the male heterotic group.

Each seed has a genetic complement from each parent. At every locus where the two inbred parents possess different alleles, the single cross hybrid is heterozygous. All F_1 seeds are genetically identical and plants are uniform.

Double-cross Hybrids and Three-way Hybrids

When the heterotic pattern is young and inbreeding depression in hybrid parents is still an issue, the use of double-cross hybrids can offer an advantage. Double-cross hybrids involve four distinct inbreds as parent lines and require two steps to create hybrid seed for farmers. Two inbreds from the female heterotic group (P1 and P2) are crossed and two inbreds from the male heterotic group (P3 and P4). Then, the F_1 seed produced by the female parents is crossed to the F_1 seed produced by the male parents. The vigor of plants serving as parents of hybrid seed production is improved due to within-heterotic-group variability (i.e. the parental lines perform better), boosting hybrid seed yield. The hybrid seed for farmers exhibits between-group heterosis.



A three-way cross is a variation on double-crossing: the hybrid seed for farmers involves a single-cross as one parent and an inbred line as the other parent. Typically, the single-cross parent is the female parent of the hybrid.

Disadvantages of Multi-cross Hybrids

Although use of double-crosses and three-way crosses can overcome effects of inbreeding depression to improve seed yields in hybrid production, there are disadvantages to be weighed as well. An extra step is required in hybrid production which demands additional time and resources in product development. Furthermore, the hybrid seed for farmers is not genetically uniform. Thus, the improved multi-cross hybrid is not expected to yield as well as a single-cross hybrid since SCA is not maximized.

Bottom line: Any new prospective cultivar must be able to be reproduced efficiently and cost-effectively!

References

Bernardo, R. 2010. Breeding for Quantitative Traits in Plants, 2nd ed. Stemma Press, Woodbury, MN.

Brar, G.S., O.N. Middleton. 1987. Genic male-sterile maize. United States Patent. Patent number: 4,654,465.

Breeze, M.L., K.S. Crowley, A. Deffenbaugh, J.R. Groat, G.D. Grothaus, D.I. Gustafson, T.O. Koyejo, T.C. Lee, D. P. Phillion, D. A. SaltmirGs, N.K. Scanlon, M.R. Starke, J. Soteres, J.E. Whitsel, M.A. Widemann. 2010. Petition for the Determination of Non-regulated Status for MON 87427 Maize with Tissue-Selective Glyphosate Tolerance Facilitating the Production of Hybrid Maize Seed. http://www.aphis.usda.gov/brs/aphisdocs/10_28101p.pdf

Chang, Z., Z. Chen, N. Wang, G. Xie, J. Lu, W. Yan, J. Zhou, X. Tang, X.W. Deng. 2016. Construction of a male sterility system for hybrid rice breeding and seed production using a nuclear male sterility gene. PNAS 113(49): 14145-14150. doi: 10.1073/pnas.1613792113

Commuri, P., N. Weber, T. Rood, R. Townsend. 2009. Petition for the Determination of Nonregulated Status for Maize 32138 SPT Maintainer Used in the Pioneer Seed Production Technology (SPT) Process. http://www.aphis.usda.gov/brs/ aphisdocs/08_33801p.pdf

Dudley, J.W. 1987. Modification of methods for identifying inbred lines useful for improving parents of elite single crosses. Crop Sci. 27: 944-947.

Dudley, J.W. 1984. A method for identifying populations containing favorable alleles not present in elite germplasm. Crop Sci. 24: 1053-1054.

Eberhardt, S.A. 1971. Regional maize diallels with US and semi-exotic varieties. Crop Sci. 11: 911-914.

Eyster, L.A. 1921. Heritable Characters of maize: VII Male sterile. J. Heredity 12: 138-141.

Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Australian J. Biol. Sci. 9: 463-493.

Hallauer, A.R., E. Lopez-Perez. 1979. Comparisons among testers for evaluating lines of corn. Proc. Corn Sorghum Ind. Res. Conf. 34: 57-75.

Matzinger, D.F. 1953. Comparison of three types of testers for the evaluation of inbred lines of corn. Agron. J. 45(10): 493-495.

Mumm, R.H., J.W. Dudley. 1994. A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. Crop Sci. 34:842-851.

Shull, G.H. 1952. Beginning of the heterosis concept. p14-48. In Heterosis. J.W. Gowen (ed). Iowa State College Press, Ames, IA.

Shull, G.H. 1909. A pure-line method in corn breeding. Rep. Am. Breeders Assoc. 5: 51-59.

Shull, G.H. 1908. The composition of a field of maize. Rep. Am. Breeders Assoc. 4: 296-301.

Sprague, G.W. 1983. Heterosis in maize: Theory and Practice, p 47-70. In R. Frankel (ed). Heterosis: Reappraisal of theory and practice. Springer-Verlag, New York, NY.

How to cite this chapter: Mumm, R.H. (2023). New Line Development and New Line Evaluation Single-cross Hybrid. In W. P. Suza, & K. R. Lamkey (Eds.), *Cultivar Development*. Iowa State University Digital Press.

Chapter 4: New Line Development and New Line Evaluation: Clonally Propagated Cultivars

Rita H. Mumm

In Chapter 2, the process and concepts of New Line Development and Evaluation were examined using soybean as an example of a pure line variety of a self-pollinated crop. Then, in Chapter 3, this process was analyzed with respect to changes needed to accommodate hybrid cultivars such as single-cross hybrid maize.

In this chapter, we will examine the changes to the process and the product pipeline that are unique to developing improved clonally propagated cultivars. The example cited is the potato (the so-called Irish potato, *Solanum tuberosum* subsp. *tuberosum*) (Fig. 1).

Fig. 1 Various kinds of potatoes. Photo courtesy of USDA ARS.

What are Clonally Propagated Crops?

Clonally propagated crops are those maintained and distributed for cultivation by asexual reproduction. Across all clonally propagated crops, different plant parts may be used as **clones**, e.g. tubers, roots, and stem cuttings. Potato varieties are maintained as tubers (Fig. 2).

Fig. 2 Potatoes are propagated as tubers. Photo courtesy of USDA.

Features of Clonally Propagated Crops

Most clonally propagated crops are cross-pollinated; many are obligate outcrossers due to self-incompatibility. This means that typically parents used in breeding crosses are highly heterozygous, and heterozygous F_1 individuals produced by crossing comprise the breeding population. Thus, the family

structure in the breeding population begins with a single plant at the F_1 stage. Actually, each F_1 plant is a potential variety!

The high degree of heterozygosity facilitates exploitation of heterosis. The level of heterozygosity is amplified by the fact that many clonally propagated crops are polyploid. For example, potato can be 2x, 3x, 4x, or 5x; Irish potato is tetraploid. Theoretically, every allele at a locus could be distinctive.

All released varieties of clonally propagated crops are **homogeneous**: i.e. varieties are non-segregating, genetically fixed, and completely uniform. At the same time, each variety may be a highly heterozygous hybrid.

Exploiting Genetic Variation

New genetic variation is generated by sexual reproduction in crossing parent lines. Once the resulting seed has been produced, no genetic changes occur. All clones created from each F_1 seed plant are genetically identical. Asexual reproduction is used to generate all subsequent generations of materials derived from the F_1 seed plants.

Thus, all genetic variation can be exploited in clonally propagated crops, i.e., additive, dominance, and epistatic variation. The covariance between clones (of different F_1 plants) equals the entire genetic variance and this variation is passed in its entirety to the next clonal generation. This represents a major advantage over breeding progeny that requires sexual reproduction to produce the next generation.

Furthermore, all testing is conducted using clones of the F_1 plants. Genotypic frequencies do not change since there is no inbreeding. Gene frequencies change only with further crossing. The genotype finally released as a new cultivar is accessible to the breeder immediately after the initial crossings!

Why Further Steps Are Needed

Theoretically, it is possible to identify the “best” clones in the first year of testing. However, there are obstacles to accomplishing this:

- Plants grown from seed do not perform comparably with plants generated from vegetative planting material. Furthermore, the plants raised from seed are typically grown in the greenhouse, which is not representative of field conditions.
- Propagation of materials for testing takes time. The amount of planting materials available at each stage is determined by the **propagation coefficient** of the crop. Potato has a relatively low propagation coefficient of approximately 10 (compared to sweet potato at 30+).
- Clonally propagated crops often involve complex product targets. Many traits are required at the farmer, processor, and consumer levels. Evaluations to facilitate selection are numerous and require significant amounts of testing materials.
- Genotype by environment interaction (**GxE**) can be very high for yield and other low heritable traits in clonally propagated crops like potato. Multiple-location, multiple-year evaluation for yield, disease/pest resistances, and quality traits are needed to satisfy product demands locally and regionally.

Examining Genotype by Environment by Management (GxExM) Systems

In addition, because all genetic variation (additive, dominance, epistatic) is captured in clonally propagated crops, GxE can be examined in depth to identify "best" genotypes for specific types of environments.

After all, *phenotype* (*P*) is the observed expression of a trait in an individual, as determined by genetic makeup and environmental factors.

$$P = f(G, E, GxE)$$

where
G = genetic effects
E = environmental effects (locations, years, seasons, etc.)
GxE = genotype x environment interaction effects

For example, phenotypic expression between clones of the same genotype grown in two different environments becomes a function of environmental variation. (Note: with only one genotype considered, GxE is not relevant or discernable.)

To explore the environmental effects and their impact on phenotype further, the particular *management* systems and practices to be employed in crop production must be detailed.

$$P = f(G, E, M, GxE, GxM, ExM, GxExM)$$

Genotype by environment by management (**GxExM**) system can be exploited to realize the full genetic potential of new cultivars in crop production.

More on Genotype by Environment by Management (GxExM)

What aspects of the management system contribute to GxExM?

Some relevant management aspects include:

- Watering regime
- Tillage regime
- Soil fertility regime
- Pest control regimes
- Plant density
- Sowing date
- Row spacing

All of these aspects have the potential to influence productivity and quality of crop production.

For potato, an important element of cultivar release is guidance to farmers on management aspects of the new cultivar in production. Thus, management practices must be addressed in the testing regime prior to new cultivar release, e.g., ridge-furrow tillage system (Fig. 3).

Fig. 3 Potatoes grown with ridge-row tillage with 36-inch row spacing. Photo courtesy of Bradley King, USDA ARS.

Complex Product Targets

Target traits for potato must meet demands of the value chain stakeholders and may include:

- High fresh-weight yield
- Yield stability
- Abiotic stress tolerances to low soil fertility, drought, heat, salinity
- Maturity
- Tuber appearance: number, shape, size, uniformity, skin color, flesh color, eye depth, lack of internal defects (e.g. hollow heart, brown center), lack of external defects (e.g., cracks, greening)
- Adaptation to local environment
- Disease resistance (e.g. late blight, early blight, bacterial wilt, potato leafroll virus, etc.)
- Pest tolerances (e.g. Potato tuber moth, green peach aphid, leafminer fly, nematodes)
- Flavor, aroma, texture
- Nutrient and anti-nutrient content (e.g. carotenoids, anthocyanins, Vitamin C, micronutrients; low levels of glycoalkaloids)
- Storage life
- Processing qualities (e.g. frying for French fries or chips, blackening after cooking)
- Starch type (altered ratio of amylopectin to amylose for industrial purposes)

See Bradshaw and Bonierbale (2010) for a comprehensive list of potential traits of interest.

Product targets are typically complex and require many types of screens to evaluate performance. Because of the fresh produce market for potato, it is critical to have farmers and consumers participate early in the selection process. However, farmer participation is best limited to selection for visual traits, especially those for which a particular threshold for performance is essential. Lowly heritable traits like yield are best selected by the breeder.

Choosing Parents

What are the priorities for choosing parents to achieve a complex breeding target in potato?

The goal of breeding crosses is to create breeding populations with high mean performance for traits of interest as well as wide genetic variability. Again, considering all traits of interest, the ideal is a good x good cross, resulting in new combinations of favorable alleles and more loci with the favorable genotype.

For autotetraploid, highly heterozygous potato, *per se* performance of parents is not necessarily a good indicator of breeding value; combining ability may be more informative. In this situation, the GCA (general combining ability) for a quantitative trait is composed of additive and additive-by-additive epistatic gene effects. For traits involving heterosis (e.g. yield), SCA (specific combining ability) may be critical to trait performance in progeny. Thus, diallel analysis may be useful in estimating breeding value (see Chapter 3 for guidelines and ALA example) and more reliable than *per se* performance in choosing parents.

Commonly, the parents include a set of 100+ clones selected in the previous year following preliminary trials for all traits. Thus, a form of recurrent selection may be practiced.

Mating Designs and Breeding Methods

Due to the prominence of self-incompatibility among clonally propagated crops, the **polycross** procedure is commonly used to create F1 progeny from multiple parents. The goal is to obtain equal genetic contribution from each parent among the progeny created.

For the polycross design to be an effective means of intercrossing multiple parents:

- Parents must be physically positioned such that there is equal chance for all possible crosses to be made.
- Parents must be flowering at the same time.
- Procedures for bulking resulting seed must facilitate equal genetic contribution of each parent.

Two experimental designs are especially useful for arranging the parents in the field to facilitate random mating.

1. The Latin Square (Fig. 4), which arranges each parent adjacent to every other parent across the replicated plantings. More than one Latin Square can be used to increase the number of replications.
2. The Randomized Complete Block (RCB) design (Fig. 5), which does not facilitate proximity of all parental pairs but does offer opportunity for more replications and can handle higher numbers of parents than the Latin Square.

Latin Square 1				
A	E	C	D	B
C	B	D	A	E
D	A	B	E	C
B	D	E	C	A
E	C	A	B	D

Latin Square 2				
A	B	D	E	C
E	C	A	B	D
B	D	E	C	A
C	A	B	D	E
D	E	C	A	B

Fig. 4 Example of a polycross arrangement for 5 parents (A through E) with 10 replications using a Latin Square design.

Randomized Complete Block (reps)									
1	2	3	4	5	6	7	8	9	10
A	D	C	B	E	D	A	C	B	B
C	E	A	C	B	C	C	E	E	A
B	A	B	D	A	A	E	D	A	E
E	B	E	A	C	B	D	B	D	D
D	C	D	E	D	E	B	A	C	C

Fig. 5 Example of a polycross arrangement for 5 parents (A through E) with 10 replications using an RCB design.

Of course, biparental crosses can also be arranged to produce progeny instead of a polycross approach.

Producing Breeding Cross Progeny

To promote flowering of difficult clones, potatoes may be grafted onto rootstocks of tomato (a related species), which may improve the amount and duration of flowering and reduce abortion of unpollinated or newly pollinated flowers (Fig. 6).

Fig. 6 Flower of the potato plant. Photo courtesy of USDA.

Once F_1 seed is generated, it can be bulked in a number of ways:

- Across the parent plants
- By parent, with an equal amount of seed from each parent bulked
- By parent, with an equal amount of seed from each replication of each parent and an equal amount of seed from each parent bulked. This approach not only ensures an equal amount of seed of each parent but a similar genetic contribution of each parental pair. Across each replication, a parent would be in physical proximity to different other plants contributing pollen.

Example Commercial Potato Cultivar Improvement Program

Consider a suitable testing regime to identify improved cultivars of a clonally propagated crop that meets a specific complex product target, using a commercial potato breeding program as an example (Table 1).

Table 1 New Line Development and New Line Evaluation in a commercial product pipeline for potato.

Season	Activity
1	Cross parent lines to generate 150-200 populations; true seed is created (~200 seeds per fruit)
2	Grow true seed in greenhouse; a single tuber is harvested from each of 30,000-40,000 seedlings. No selection except perhaps against disease-infected plants.
3	Grow tubers in single row plots with 3-5 plants in unreplicated trial at 1-2 locations. Visually select against undesirable traits: e.g. plant type, excessively late maturity, unacceptable tuber type.
4	Grow >5,000 selected clones in 2-3 row plots at ≥2 locations (one may be a disease screening site). 1. Select visually for highly heritable traits (e.g. color, traits with threshold requirements) 2. Select visually (or with ELISA) for disease/pest resistances with farmer participation 3. Select for yield and quality traits
5	Grow 100-200 selected clones in 3-5 row plots at multiple locations including marginal or high stress environment. Evaluate yield and all agronomic traits including yield stability as well as post-harvest traits and quality traits; farmer participation in selection for visual traits.
6	Evaluate 50-100 selected clones for yield, agronomic, quality, post-harvest traits in 5-row plots at multiple locations; farmer participation, particularly for flavor and post-harvest; best clones coded.
7	Advanced replicated yield trials by maturity group in many locations.
8	1. Advanced replicated yield trials by maturity group in many locations. Disease and quality screening continues. 2. Agronomic trials to develop crop management specifications. 3. S-clone propagation ("seed clones" produced as a source of planting material) and *in vitro* maintenance of disease-free true seed plantlets.
9	On-farm trials and large-scale grower trials; evaluate consumer acceptance as fresh or processed product.
10	Variety registration
11	Release 0-2 new varieties

Salient Features

We observe the following with respect to the Example Commercial Potato Cultivar Improvement Program:

- A large number of crosses are made, with at least 200 progeny per cross.
- The "true seed" is grown in the greenhouse to establish the F_1 plants that represent the families in the breeding crosses.
- The first selection focuses on elimination of undesired genotypes (not selection for top performance) with moderately high selection intensity.
- Early testing stages focus on discarding undesirables, selecting for high-heritability traits and those that can be scored visually; advanced testing stages focus on selection for yield and other lowly heritable traits, while continuing selection for other key traits like disease resistance.
- Plot sizes increase and testing locations expand as testing advances with selected clones.

- Farmers are engaged early in participatory breeding and throughout the selection process, especially to provide input on visually-scored traits, flavor, and post-harvest characteristics.
- Selection approaches may include independent culling (if a clone does not meet selection thresholds for one trait, it is discarded in other tests preferably before harvest); index selection facilitates selection in a complex product target for multiple traits that may be negatively correlated with yield.
- The selected clones from Season 6 are "coded" and progressed to advanced testing at numerous locations.
- Yield and other lowly heritable traits influenced by GxE are evaluated for multiple years in multiple locations.
- **S-clones** of selections advanced through regional trials are propagated in concert with advanced testing to prepare for release and distribution of potential new varieties. A disease-free source for S-clone propagation must be maintained.
- Commercial grower trials provide evaluation of bulk handling and consumer/processor acceptance as a fresh or processed product.
- Plant materials for testing can originate with the previous year's trials unless there is a disease build-up.

How does this breeding and testing regime differ from one focused on development of improved single-cross corn hybrids?

More About S-clones

With clonally propagated crops, the source of planting materials for cultivar production by farmers is not seed. It is the clonally propagated plant part (e.g., tubers) of the new, improved variety. For potato, these source planting materials are tubers referred to as S-clones or "seed clones" (Fig. 7). Because sexual reproduction has no place in maintaining the S-clones, there is no risk of contaminating the new, improved variety due to pollen migration. The only genetic changes to the variety would occur through mutation, which is rare. However, disease transmission is a risk as many diseases can be spread through vegetative plant parts.

Fig. 7 Uprooted potato tubers from the soil. Photo courtesy of Belize Ag Report.

A newly developed variety can be lost without adequate safeguarding from disease infection. After all, a clone developed from two heterozygous parents cannot be reproduced by crossing the parents again. Virus diseases are particularly threatening since viruses cannot be controlled chemically and are often spread by insect vectors. Thus, to have a source of healthy base material, true seed plantlets may be maintained *in vitro* under disease-free conditions until final selections are made from each batch. This virus-free starter material can be used to replace S-clones of new, improved varieties that are overcome in field environments. It is the breeder's responsibility to provide disease-free starter materials for S-clone

propagation as part of the hand-off to Supply Chain (the core function in the process pipeline charged with producing materials for distribution of new varieties to farmers).

Other Clonally Propagated Crops

Fig. 8 Cassava is the basic staple food for more than 40% of the people in Africa. Photo courtesy of Food and Agriculture Organization of the United Nations.

Besides potato, other crops that are typically maintained and distributed for cultivation by asexual reproduction include the below (some shown in Figs. 8 and 9):

- Cassava
- Sweet potato
- Yam
- Taro
- Banana
- Plantain
- Sugar cane
- Apple
- Strawberry
- Grapes

Fig. 9 Left to right: banana, sweet potato (photos by Scott Bauer); grapes and apple. Photos courtesy of USDA.

References

Bradshaw, J.E., and M. Bonierbale. 2010. Potatoes. In J.E. Bradshaw (ed). Handbook of Plant Breeding: Root and Tuber Crops. Springer Science+Business Media, LLC, New York.

How to cite this chapter: Mumm, R.H. (2023). New Line Development and New Line Evaluation Clonally Propagated Cultivars. In W. P. Suza, & K. R. Lamkey (Eds.), *Cultivar Development*. Iowa State University Digital Press.

Chapter 5: Value-added Trait Integration

Rita H. Mumm

The **cultivar development process** is comprised of four main functions (as discussed in Chapter 1. Earlier chapters have focused on New Line Development and New Line Evaluation. This chapter will focus on the next step in the product pipeline: value-added trait integration, or "Trait Integration" for short. This core function serves to incorporate high-demand traits to further improve elite cultivars (Fig. 1).

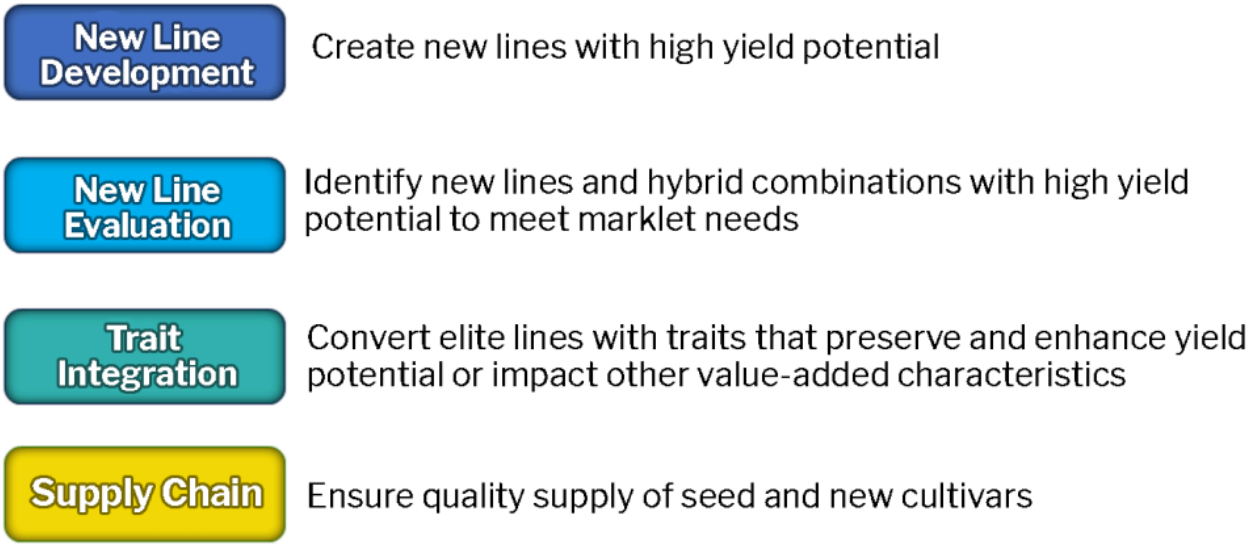


Fig. 1 The four core functions in the product pipeline.

What is a 'Value-added' Trait?

A **value-added trait** (VAT) is a special trait that represents a novel or uncommon characteristic of value to farmers, end-users or consumers *when incorporated into an elite cultivar*.

Fig. 2 Lettuce with potyvirus. Photo courtesy of the Richard W. Michelmore Lab, University of California, Davis.

An example in lettuce is resistance to the potyvirus, Lettuce Mosaic Virus (LMV), which causes devastation in lettuce production worldwide and can result in severe crop losses (Fig. 2).

Symptoms vary by variety and timing of infection, yet generally involve deformed heads and discolored leaves. A single, recessive gene, *mo1*, confers resistance/tolerance.

Think of a VAT as "frosting on the cake!" Typically, it is a "must-have" characteristic for one or more stakeholder groups in the crop value chain, and sometimes such a trait commands a premium price depending on the economics of its deployment.

More About Value-Added Traits

Typically, VATs involve no more than five genetic factors and often are conferred by single genes. Generally, these traits are not common in the gene pool. Some have designated tradenames.

VATs include traits of all types, for example:

- Disease resistance
- Insect resistance
- Herbicide tolerance
- Abiotic stress (e.g., drought tolerance, low fertility tolerance)
- Yield/productivity enhancement
- Nutritional enhancement, e.g., more lysine (corn), healthy oil profile (soybean), reduced lignin content (alfalfa)
- Consumer or end-user preferences, e.g., non-browning (apple), reduced black spot bruising (potato), fruit ripening (tomato)
- Male sterility

- Flower color (rose)

Developing VATs

VATs may be developed through different techniques such as:

- **Mutagenesis**.
- **QTL** mapping.
- Transformation (e.g., virus resistant papaya)
- Gene editing (e.g., Sulfonylurea herbicide tolerant canola)

Discussed below are some crop improvement examples through the use of these techniques.

VATs Based on Mutant Genes

Mutagenesis can generate new alleles, some of which may have novel utility in a crop plant.

An example is imidazolinone-herbicide tolerant rice Clearfield® rice developed by BASF. It confers tolerance to a class of herbicides which inhibit amino acid synthesis by inhibiting the acetolactate synthase (ALS) enzyme. The ALS mutant allele was created through chemical exposure to ethylmethanesulfonate (EMS), a mutagen.

Most mutant alleles are recessive, which has implications for product development. For example, for threshold expression of the trait in a hybrid cultivar, two copies of the mutant allele would be required. This demands that both inbred parents of the hybrid carry the allele.

TILLING

TILLING (**T**argeting **I**nduced **L**ocal **L**esions **IN** **G**enomes) provides a practical path to discovery of single-nucleotide mutations. TILLING is a **reverse genetics** approach; that is, gene function is determined by analyzing the phenotypic effects of specific engineered gene sequences. Reverse genetics seeks to learn what phenotypes arise as a result of particular genetic sequences. In contrast, **forward genetics** seeks to understand the DNA sequence that gives rise to a particular phenotype.

Approaches to determining gene function:

- **Forward genetics**: Determine the DNA sequence that code for the known phenotype.
- **Reverse genetics**: Determine the phenotype that will be produced from DNA sequence.

Steps in TILLING

Till et al. (2004) lists steps to finding changes in specific genes of interest (Fig. 4):

1. Identify specific genes of interest; target DNA regions to screen
2. Create primers to PCR target regions
3. Apply mutagen (e.g. ethylmethanesulfonate [EMS]) to pollen or seed
4. Screen for mutations in **M_1** (if from pollen) or **M_2** generation (if from seed)
5. Identify, sequence, and phenotype mutant individuals.

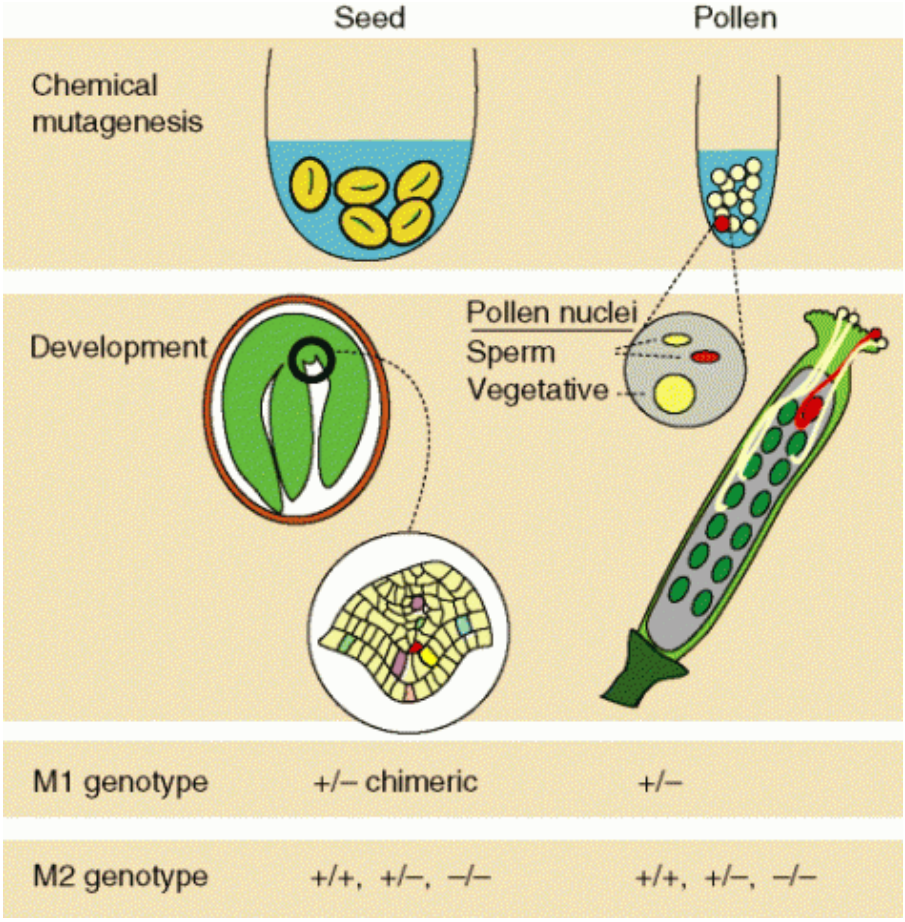


Fig. 3 Genetic constitution of M_1 and M_2 resulting from mutagenesis of seed and pollen. Image from Comai and Henikoff, 2006.

The process has been improved and continues to be optimized (Fig. 4).

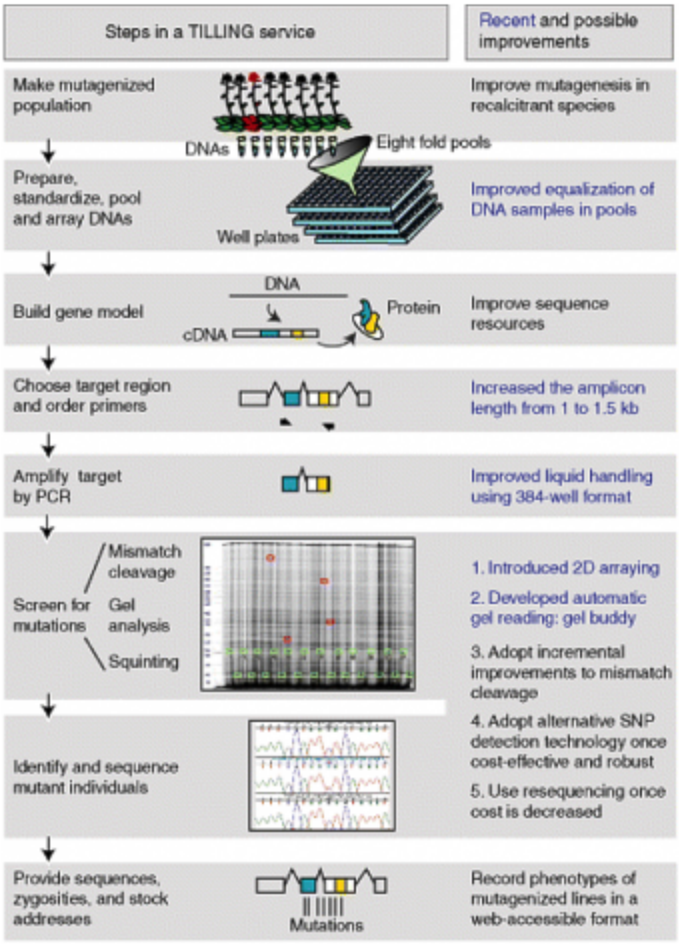


Fig. 5 Steps in an improved TILLING process. Image from Comai and Henikoff, 2006.

VATs Based on QTL

Some VATs are the result of QTL mapping in which native genes have been tagged by molecular markers. Examples include:

- **Optimum AQUAmax®** corn hybrids by DuPont Pioneer (now Corteva Agriscience) which are drought tolerant. Chromosomal segments identified through marker-assisted breeding to withstand water stress are assembled and deployed. (Knowledge of specific genes [quantity, function, modes of action] is not necessary.)
- **AgrisureArtesian** one of **Syngenta Artesian™** drought tolerant corn hybrids. To develop this VAT, candidate genes associated with water-deficiency stress tolerance were identified through genomics. This led to the discovery of maize genes that protect the plant from drought stress in several different ways, responding to water-deficiency stress with multiple modes of action and at various stages of growth through the corn lifecycle. The genes were validated and pyramided in elite hybrids to deploy the trait (Fig. 5).

Fig. 5 Artesian hybrid (top right, bottom left) vs. non-Artesian hybrid (top left, bottom right). Audrain County, Missouri, USA, 2012. Photo courtesy of Syngenta.

VATs Based on Transformation

VATs developed through transformation (also referred to as **genetically modified** or **GM traits**) have primarily involved protection of the genetic yield potential of elite cultivars against weed, insect, and disease pests. By and large, these VATs represent novel traits that are either not accessible in the current host species gene pool or which have resulted from the modification of host species genes for expression at higher thresholds, in specific tissues, at certain timing in the host species life cycle, or for **silenced** expression.

Transformation produces a transgenic **event,** which becomes the original source of the VAT. An event is defined by the unique DNA sequence inserted in the host genome through transformation and the precise point of insertion. Each event originates as a single plant (referred to as a T_0 plant) which was regenerated from a single transformed cell.

Examples of Transgenic VATs

Examples of VATs developed through transformation among others include insect resistant corn and virus resistant papaya:

- YieldGard® developed by Monsanto Company, confers resistance to Lepidopteran insect pests that feed on leaves, stalks, and ears in corn.
- Papaya resistant to Papaya Ringspot Virus (PRSV) was introduced in Hawaii USA in 1998 and was rapidly adopted by Hawaiian growers. Resistance, developed by then Cornell University researcher Dennis Gonsalves, is conferred through gene silencing in response to a virus coat protein gene fragment, Event 55-1. The most commonly grown cultivars are Rainbow, Sunrise/Sunup, Kapoho (Solo), and Kamiya/Laie Gold.

Rapid Adoption of GM Crops

From 1995, when GM crops were first commercialized, through 2015, the land area around the globe planted to GM crops increased from 1.7 million hectares to 179.7 million hectares, making this technology the most rapidly adopted in modern agriculture (Fig. 6).

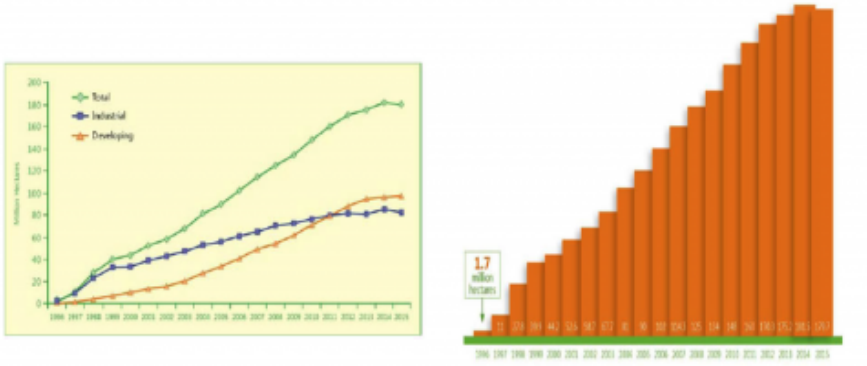


Fig. 6 Global area of GM crops, 1996 to 2015, and in industrial and developing countries (in million hectares). Image from Clive James, 2015.

The primary GM crops globally are soybean, maize, cotton, and canola (Fig. 7).

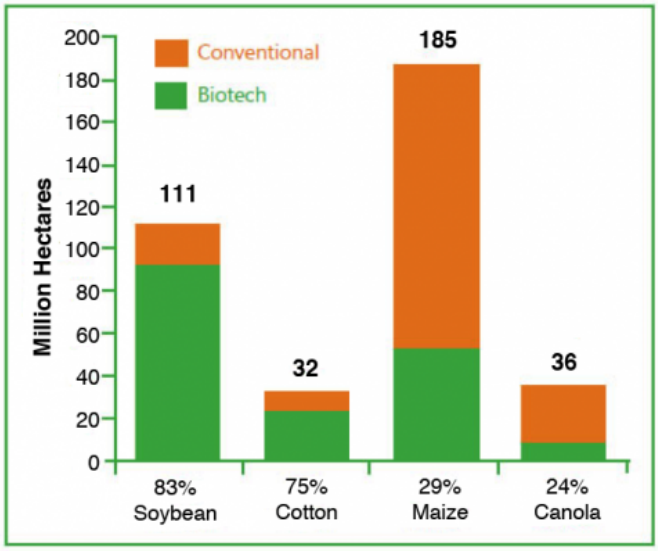


Fig. 7 GM crop area as a percentage of the global area of principal crops, 2015 (in million hectares). Data from FAO, 2013. Image from ISAAA.

GM Crop Adoption in Africa

GM crops are widely grown and used throughout the world. In Africa, three countries have approved at least one type of GM crop: Burkina Faso (cotton), Sudan (cotton), and South Africa (maize, soybean, cotton). Others have approved research field trials: Cameroon, Egypt, Ghana, Kenya, Malawi, Nigeria, and Uganda (Fig. 8).

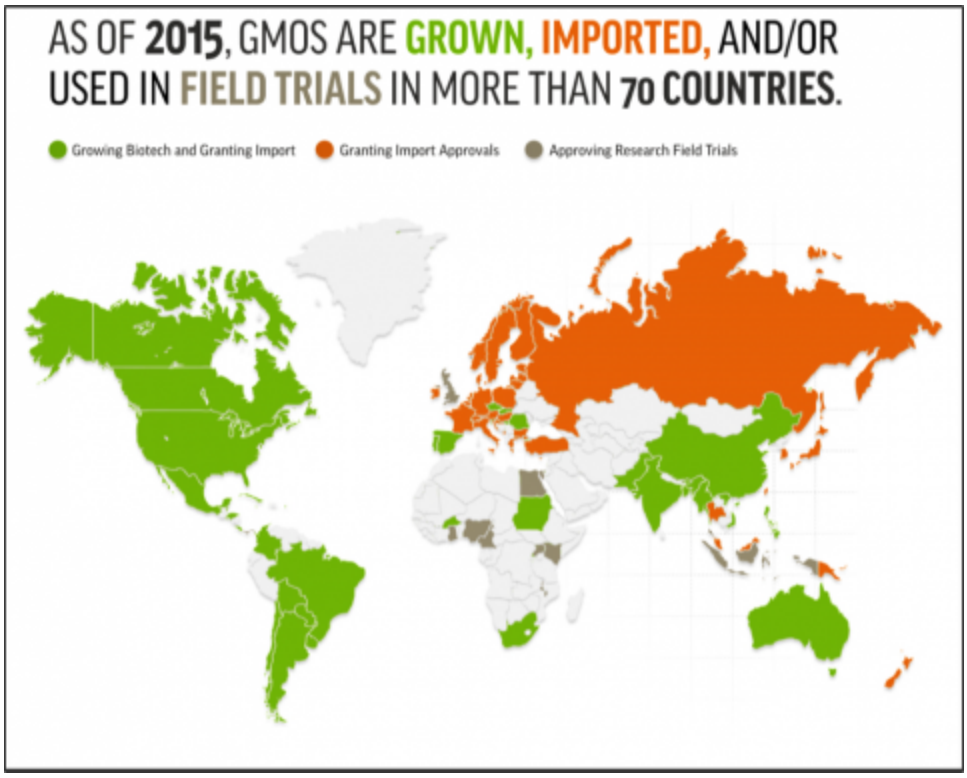


Fig. 8 Global GMO status in 2015. Image from GMO Answers.

For more on the status of biotech crops in Africa, check out this video by ISAAA.

Stacked GM Traits

Not only have GM traits been rapidly adopted as VATs, the proportion of cultivars with combined, or **stacked**,GM traits has escalated quickly. Crops containing more than one GM trait were grown on nearly 60 million hectares globally in 2015 (Fig. 9).

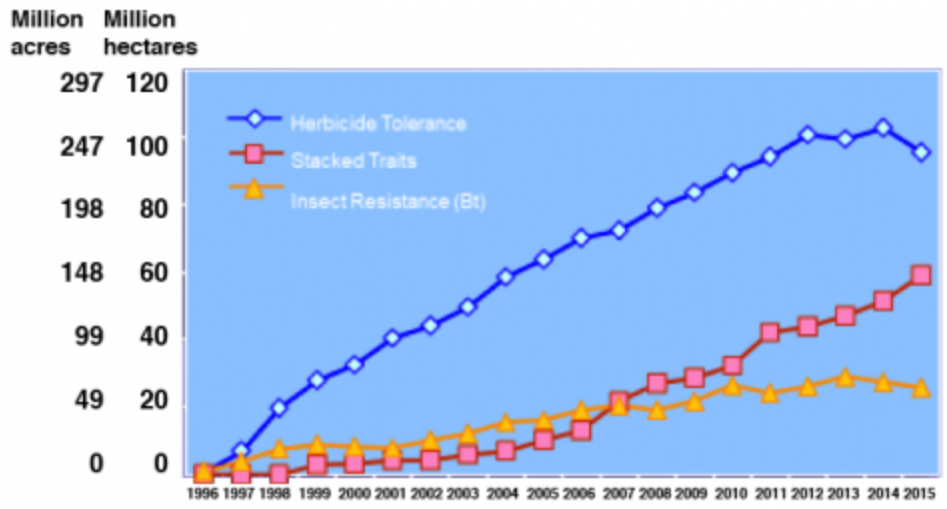


Fig. 9 Global area of biotech crops, 1996-2015, by trait. Image from Clive James, 2015.

SmartStax® Example

Not only do the stacks represent multiple value-added characteristics, some represent multiple genes (or events) for the same characteristic.

Let's look at an example that demonstrates both scenarios: SmartStax® corn.

SmartStax® contains 8 genes conferring VATs:

- 3 genes for *above-ground insect resistance to Lepidopteran species*
- 3 genes for *below-ground insect resistance to Corn Rootworm*
- 2 genes for herbicide tolerance: one conferring *tolerance to glyphosate herbicide* and one conferring *tolerance to glufosinate herbicide*.

Thus, there are 4 VATs present in SmartStax®.

The sets of multiple genes for above- and below-ground insect resistance make it difficult for the insect species to overcome, whereas single genes are relatively easily overcome when enough selection pressure is applied. This pyramiding of genes conferring a given trait is a "best practice" as it is a key strategy for managing the development of resistance in pests and prolonging the utility of an architected solution (Fig. 10).

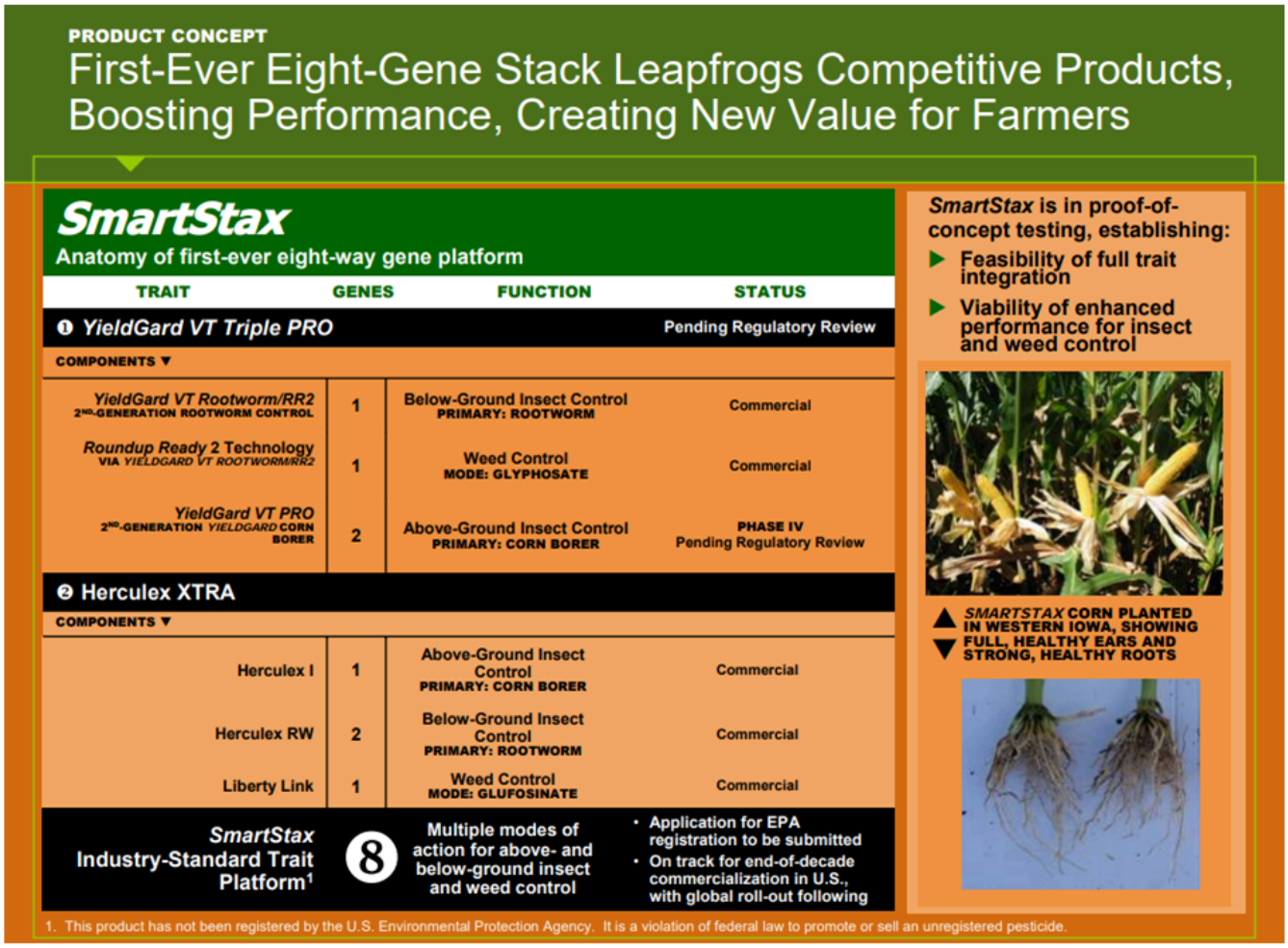


Fig. 10 Description of the elements comprising the SmartStax product according to information released Sept 14, 2007. All events included were subsequently approved for cultivation in the USA. Source Monsanto-DowAgroSciences, 2014.

Also note that there are four events involved in delivering the 8 genes: MON88017, MON89034, TC6275, and DAS59122-7 for this product. Each event includes more than one gene, a condition referred to as "**molecular stacking**".

GM Events Require Governmental Approvals

VATs created through genetic engineering are subject to government regulation on a country-by-country basis. Before approving cultivation or import of GM crops, governmental bodies generally review the issues related to:

- Problem addressed and effectiveness of the proposed solution.
- Food safety.
- Safety to the environment and biosphere.

The adoption of transgenic VATs in some countries is hampered by the lack of appropriate governmental bodies to review and authorize use.

Breeding strategies and transport of seed must take account of government regulations and comply with containment policies.

VATs Based on Gene Editing

Gene editing (or genome editing) comprises a range of molecular techniques that facilitate targeted changes to be made in the genome. It involves the use of certain nucleases to make precise cuts in the DNA, which then can become the sites of base pair substitutions, DNA deletions or DNA insertions, harnessing the organism's own DNA repair system. Gene-editing techniques are typically named for the type of nuclease used.

For example, CRISPR-Cas9 involves:

- **C**lustered
- **R**egularly
- **I**nterspaced
- **S**hort
- **P**alindromic
- **R**epeats

coupled with a programmable nuclease derived from bacteria (Cas9).

For more details on genome editing, see the following YouTube videos:

- Genome Editing with CRISPR-Cas9:
- What is CRISPR?

Examples of VATs Created Through Gene Editing

Examples of VATs created through gene editing include:

SU Canola™ – was developed by Cibus to provide a non-transgenic option to canola growers for weed control by conferring tolerance to sulfonylurea herbicides and is promoted for use with Draft™ herbicide from Rotam.

Fig. 11 Mushrooms. Photo courtesy of Yinong Yang, Pennsylvania State University.

Non-browning mushrooms – were developed by Professor Yinong Yang at Pennsylvania State University using the CRISPR-Cas9 system. Several genomic deletions facilitated through gene editing have resulted in a product with longer shelf life and less post-harvest loss (Fig. 11). Commercialization is anticipated in the near future.

Maximizing the Use of VATs

You may have noticed the VATs that have been commercialized are branded (tradename, logo, etc.) to drive customer/consumer recognition and loyalty. Because VATs are in high demand and often have been developed at substantial investment, there is a desire to make the most of their use. From a breeding standpoint, this typically involves integrating each VAT into a wide array of elite cultivars to maximize market penetration.

Note that although stacking has been illustrated herein using a GM trait example, any type of VATs may be pyramided in a cultivar. Mixtures of types of combined VATs are not uncommon; many times the same customer/consumer base demanding one type of VAT may also demand others for maximum value.

Next, we will discuss ways to facilitate Trait Integration of single or multiple VATs through breeding.

Trait Integration

How does Trait Integration fit into an overall breeding program?

New elite lines produced through New Line Development and New Line Evaluation are **converted** for the VAT of interest through backcross breeding.

Elite Cultivar ➡ backcross breeding ➡ Elite *Value-added Trait* Cultivar

As a result, the breeder aims to recover the *complete and unaltered* agronomic package represented by the elite line *plus* the desired expression of *all* the genes or events introduced in the elite VAT cultivar.

Backcross Conversion Example

The general formula for integrating a single **genetic factor** (e.g., gene) via backcross (BC) conversion requires a cross between the trait donor (non-recurrent parent, NRP) and the elite line to be converted (recurrent parent, RP). Ideally, in each generation, only those individuals who inherit the desired gene are used as parents to create the next generation. Several generations of backcrossing to the RP are needed to recover the vast majority of the RP germplasm.

Once this is accomplished, individuals in the final backcross generation are self-pollinated to "**fix**" the introgressed gene in homozygous state. Progeny testing may be performed to identify non-segregating lines to bulk and label as the converted cultivar. In all, 11 or more generations may be required to complete the introgression of the desired gene (Table 1).

According to genetic theory, after six backcrosses the progeny are, on average, greater than 99% genetically similar to the RP.

Table 1 Breeding plan depicting the general formula for backcross conversion and recovery of the RP germplasm.

Generation	Plant	Activity	Produce
1	NRP, RP	Cross NRP to RP	F_1
2	F_1	Select for desired gene; BC to RP	BC_1
3	BC_1	Select for desired gene; BC to RP	BC_2
4	BC_2	Select for desired gene; BC to RP	BC_3
5	BC_3	Select for desired gene; BC to RP	BC_4
6	BC_4	Select for desired gene; BC to RP	BC_5
7	BC_5	Select for desired gene; BC to RP	BC_6
8	BC_6	Self	BC_6S_1
9	BC_6S_1	Self	BC_6S_2
10	BC_6S_2	Identify line phenotypically similar to RP that stably expresses introduced trait	BC_6S_3
11	BC_6S_3, tester	Create seed for testing performance equivalency; testcross (if RP is a hybrid parent) and/or self	BC_6S_3 TC BC_6S_4

Recovery of RP Germplasm

Genetic theory indicates with each cross to the RP, the average percentage of NRP germplasm is reduced by half, assuming no selection. The general formula for calculating the estimated mean percentage of RP germplasm among backcross progeny in a given generation is:

$$\%RP(n) = 1 - (\frac{1}{2})^{n+1}$$

where:

%RP = percentage of recurrent parent germplasm recovered
n = backcross generation number

In the BC_6 generation, progeny on average are 99.22% genetically similar to the RP (Fig. 12).

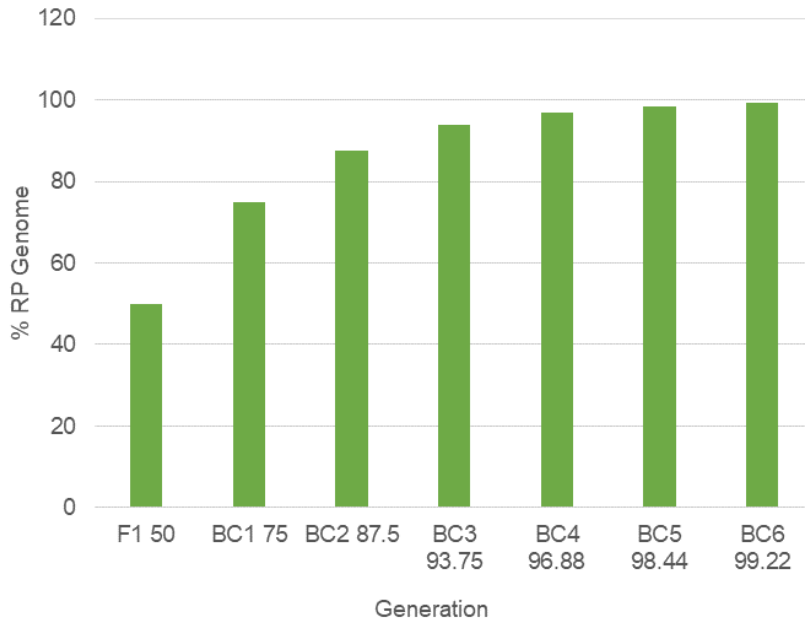


Fig. 12 Expected percentage of RP germplasm recovered with each cross to the RP. Image from data in Table 28-2; Fehr, 1987.

Backcross Population Distribution of %RP

However, all individuals in a given backcross generation do not contain the same proportion of RP germplasm. Each backcross generation represents a distribution of the percentage of RP germplasm. For example, the BC_1 generation percentage of RP germplasm is normally distributed with a mean of 75% and a standard deviation of 4.86% (Fig. 13).

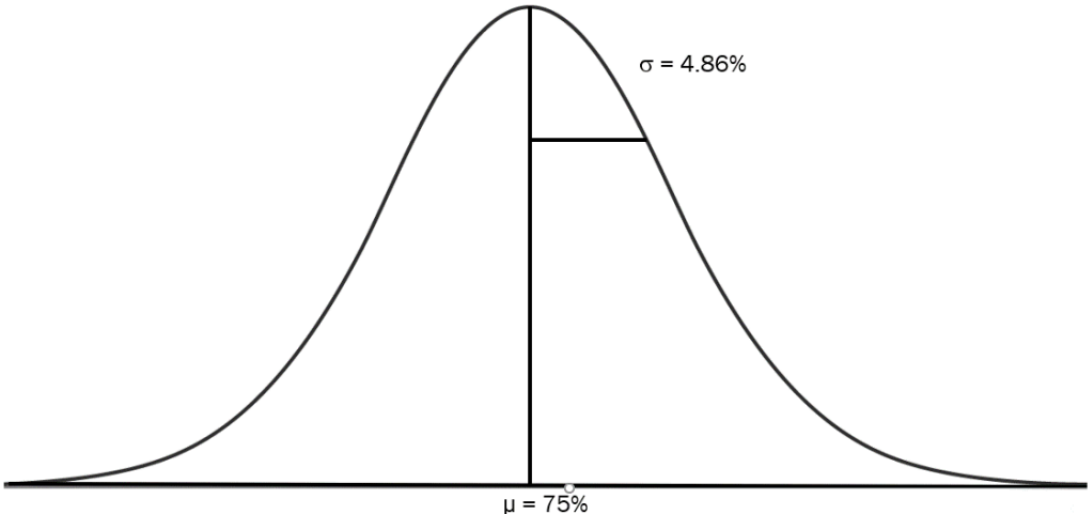


Fig. 13 Distribution of a BC1 population.

Distribution of Residual NRP Germplasm

However, in practice, with selection for the desired gene, more NRP germplasm remains and its distribution across the genome is not uniform.

Based on a 1788 centimorgan (cM) map of maize spanning 10 chromosomes, the mean amount of residual NRP germplasm was estimated using computer simulation (Peng et al. 2014). Note that the relationship between cM NRP and %RP is given by:

$$\%\mathrm{RP} = \left[1 - \frac{1}{2}\left(\frac{\mathrm{cM\ NRP}}{\mathrm{total\ NRP}}\right)\right] \times 100$$

Thus, 120 cM of NRP germplasm in a genome of 1788 cM equates to approximately 96.6% RP.

For each BC generation, 1000 progeny were screened to identify those with the introgressed gene, with approximately half selected (single gene, Mendelian segregation). The mean NRP germplasm was estimated with these 500 individuals. In backcross generations BC_1 through BC_6, the mean residual NRP germplasm (in cM) was measured across the whole genome, on chromosomes not carrying the desired gene, on the chromosome carrying the desired gene, and in the region comprising 10 cM to each side of the desired gene (20 cM flanking region).

Although NRP germplasm on non-carrier chromosomes decreases incrementally with each backcross generation, the residual NRP germplasm on the carrier chromosome decreases at a much less rapid rate. Furthermore, the amount of residual NRP germplasm in the flanking region remains somewhat stagnant through 6 backcrosses; selection for the desired gene is pulling along linked DNA from the donor parent (Table 2).

Table 2 Mean amount of residual NRP germplasm in maize backcross generations as shown through computer simulation. Data from Peng et al. 2014.

Total Genome NRP (cM)	**Non-Carrier** Chromosomes NRP (cM)	**Carrier** Chromosome NRP (cM)	**Flanking Region on the** Carrier Chromosome NRP (cM)	n/a
BC_1	1398.79	1240.62	158.17	19.53
BC_2	973.90	846.89	127.01	18.88
BC_3	681.45	578.13	103.32	18.26
BC_4	480.31	395.15	85.16	17.66
BC_5	343.20	271.49	71.71	17.12
BC_6	248.19	187.23	60.96	16.58

Linkage Drag

DNA in the chromosomal regions flanking a VAT is under pressure when backcross progeny with the VAT are identified and selected; recombination is reduced due to linkage with the VAT. The residual NRP in the chromosomal proximity with the desired gene is referred to as **linkage drag** (Fig. 14).

The effect of this linked DNA depends on the genetic contents of these chromosomal segments. Linkage drag can potentially affect performance for key traits like yield and stress tolerance.

Why?

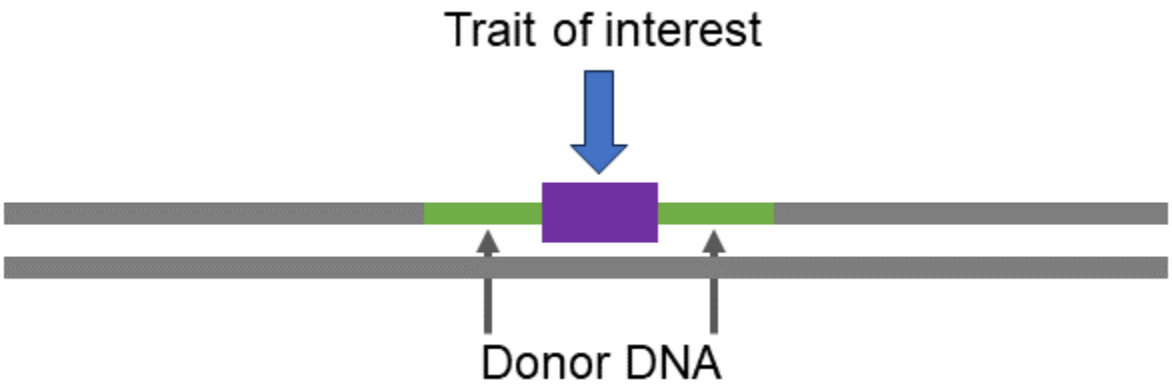


Fig. 14 Linkage drag due to residual NRP DNA in the chromosomal proximity with the desired gene.

This NRP germplasm may contain deleterious genes, especially if the trait donor is non-elite. If the introgressed genetic factor is a result of genetic engineering, this DNA from the NRP may have mutated during tissue culture, causing what is referred to as **somaclonal variation.** In the case of hybrid cultivars, if the trait donor is from the heterotic group opposite the RP, this DNA may decrease potential heterosis in the converted hybrid.

Linkage drag has the potential to interfere with the recovery of the performance of the RP and the goal to recover the *complete and unaltered* agronomic package represented by the elite line targeted for conversion! The impact of linkage drag can be magnified with the accumulated effect of multiple introgressed genes.

How can linkage drag be eliminated?

Identifying Homozygotes for the Desired Gene

Finally, once backcrossing has attained the desired level of RP germplasm recovery, the expression of the desired gene is stabilized through self-pollination to achieve homozygosity.

Since S_1 plants will be segregating for the desired gene, continued self-pollination may be required to produce materials for progeny testing.

Efficiency in Trait Integration

Since Trait Integration requires additional time beyond New Line Development and New Line Evaluation and additionally presents some risk in recovering the full set of performance attributes of the line targeted for conversion, there may be some opportunities to utilize technology in an efficient manner to reduce these negatives. Specifically, the application of technology to accelerate the conversion process and to reduce the risk of failure to recover performance equivalency to the RP would represent an enormous improvement in efficiency.

Strategically, technology addressing the following issues would make Trait Integration more efficient and effective, while increasing the rate of genetic gain:

- Selection accuracy
  - Effective screen for desired gene/trait.
  - Identification based on genotype (vs. phenotype).
- Acceleration of the process
  - Faster cycling of generations.
  - Quicker recovery of RP germplasm in backcrossing.
  - Fewer generations for gene/trait fixation.
- Quality outcomes
  - Eliminating linkage drag.
  - Ensuring a high degree of recovery of %RP.

Use of Molecular Markers for Greater Efficiency in Trait Integration

Molecular markers are a great fit for technology to meet the need for selection accuracy, acceleration, and quality outcomes. Molecular markers provide knowledge of genotypes among backcross progeny. The use of this technology has great advantages over trait integration (TI) based on phenotype alone (Table 3).

Table 3 Increase in recurrent parent germplasm, efficiency, and speed of trait fixation through use of molecular markers and marker-assisted backcrossing.

Molecular markers can be used to:	n/a	**Conduct marker-assisted backcrossing (MABC) to:**
Identify individuals that have inherited the desired gene	to	Be more efficient than phenotypic selection
Select for recovery of RP germplasm in backcross generation	to	Trim ≥3 generations from the TI process to save 1-2 years in the development of VAT hybrids
Select against linkage drag	to	Select against linkage drag to increase the probability of obtaining acceptable (i.e. quality) conversion
Identify homozygotes for the VAT gene(s) in the selfing generations	to	Reduce the number of generations to trait fixation

MABC Selection Schemes

With marker(s) for the desired gene, a set of markers that provides coverage of the genome, and a set of dense markers (one per cM in the 20 cM region flanking the desired gene), it is possible to implement (Fig. 15):

1. Selection for the target gene/event (selection for those individuals that carry the desired genetic factor).
2. Selection against linkage drag (selection for individuals with crossovers in the flanking region).
3. Selection for RP recovery (selection for those individuals in the upper tail of the backcross population distribution).

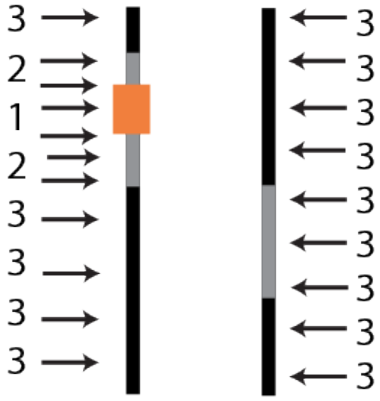


Fig. 15 Target gene (orange box), set of markers for genome coverage, and dense markers for MABC.

Designing an Efficient Trait Integration Process

Although we have been discussing the principles of single gene integration, in real life the integration typically involves multiple VATs or multiple genetic factors. Thus, pyramiding of all the genetic factors involved is necessary.

Multiple Trait Integration involves four steps (Fig. 16):

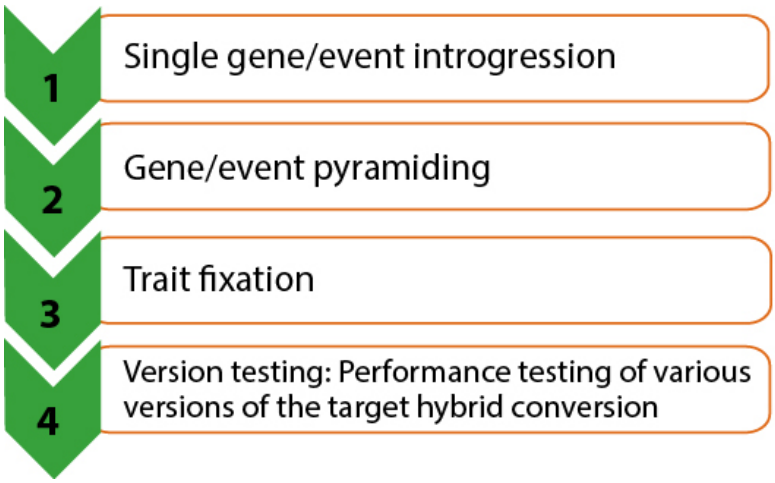


Fig. 16 Steps in multiple trait integration scheme.

Computer simulation has value in designing a product pipeline to improve the process of Trait Integration, especially since there are a considerable number of options when converting a particular cultivar for multiple genetic factors. Computer simulation offers the opportunity to identify options that can garner efficiency in terms of speed to market, rate of gain, conservation of resources, etc.

An example of using computer simulation to drive a strategic approach for multiple Trait Integration is given through a study conducted in the Mumm Lab at the University of Illinois.

Efficiencies Revealed Through Computer Simulation

A study was conducted to explore the limits of Multiple Trait Integration and identify strategies that could result in recovery of equivalent performance in a corn hybrid converted for 15 events (Peng et. al. 2014a, 2014b; Sun and Mumm 2015). Conversion for 15 genetic factors was considered to be pushing the boundaries of what could be accomplished with backcross conversion since traditionally a limit of about 5 stacked genes was accepted practice.

The breeding scenario assumed:

- Some events were required in the male parent of the hybrid so a decision was made to balance the 15 events between the female and male RPs: 8 events would be introgressed into female RP and 7 events into male RP.
- All events are new so conversions for each event are needed.
- Residual NRP germplasm in the 20 cM region flanking the introgressed event (FR NRP) will be unalterable after Step 1 (once pyramiding is initiated).
- Population size of 400 each generation, with final selection of the top four plants (selected proportion = 0.01).

Since 120 cM of NRP germplasm (~ 6.7% NRP) is the maximal amount of residual NRP germplasm consistent with recapturing target hybrid performance, *a goal for single event conversions was set at ≤8 cM Total NRP including ~1 cM of FR NRP.*

Optimization Criteria

To identify optimal breeding strategies for Step 1 in the Multiple Trait Integration process with the aim to minimize accumulation of linkage drag in the converted target hybrid and maximize efficiencies, several criteria by which to assess efficiency were defined:

1. Total NRP (cM).
2. FR NRP (cM).
3. Time (expressed in generations).
4. Marker data points (MDP).
5. Population size (N).

The greatest priority is given to Numbers 1 and 2 since these criteria determine the effectiveness of the integration process outcome, that is, the ability to recover equivalent performance. Without this, all resource expenditures and time investment in Trait Integration to further improve the elite hybrid is meaningless. Number 3 impacts genetic gain in a major way. And Numbers 4 and 5 determine resources in terms of budget, manpower, and facilities (e.g., genotyping costs, labor, greenhouse, field space).

Findings of Computer Simulation Study

The goal of the computer study was to identify an optimal strategy for selection for the particular event (E), selection against linkage drag (LD), and selection for percent RP germplasm (RP).

Preliminary investigation showed, based on selection for E+RP, that Total NRP would be reduced to about 12 cM after 10 generations of backcrossing, yet most of the residual NRP germplasm would reside in the flanking region (FR).

On the other hand, if selection was based on E+LD, the Total NRP remained high but FR NRP was reduced to ~1 cM by BC_3 or BC_4 and remained largely static after that. It seemed highly unlikely that FR NRP would be reduced much beyond 1 cM, even with extensive additional backcrossing (Fig. 17).

Reference Population with E+RP selection
(Pop size=1000/generation, Repeat=1000)

Generation	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10
Total NRP (cM)	452.83	100.44	21.09	14.93	14.29	13.75	13.31	12.89	12.5	12.07
FR NRP (cM)	17.99	15.59	12.54	11.97	11.73	11.55	11.36	11.17	10.96	10.72

Reference Population with E+LD selection
(Pop size=1000/generation, Repeat=1000)

Generation	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10
Total NRP (cM)	903.18	424.45	312.84	236.3	182.44	140.31	108.29	84.23	66.04	51.74
FR NRP (cM)	9.65	1.84	1.08	0.99	0.96	0.95	0.94	0.93	0.92	0.91

Fig. 17 Total NRP and FR NRP with 2 selection strategies across 10 BC generations. Data from Peng et al., 2014a.

Additional Simulation Results

All combinations of selection for E, LD, and RP were considered with the goal of reducing Total NRP to ≤8 cM and FR NRP to ~1 cM.

The objective was nearly met with 5 generations of selection, the first 3 generations for E+LD followed by 2 generations of E+RP. Total NRP was reduced to 7.86 cM. However, the FR NRP was at 1.68 cM — more work was needed (Table 4).

Table 4 Breeding strategies comparison. Data from Peng et al., 2014a with N=400, Repeat=1000.

Selection Methods	MAS Gen.	BC_1	BC_2	BC_3	BC_4	BC_5	Total Non-RP (cM)	FR Non-RP (cM)	MDP (K)	Ntotal
Three-Stage	5 MAS	E+LD+RP	E+LD+RP	E+LD+RP	E+LD+RP	E+LD+RP	17.85	1.42	222	2000
	4 MAS	E	E+LD+RP	E+LD+RP	E+LD+RP	E+LD+RP	31.02	1.80	178	1608
	3 MAS	E	E	E+LD+RP	E+LD+RP	E+LD+RP	47.70	2.67	134	1216
	2 MAS	E	E	E	E+LD+RP	E+LD+RP	75.68	5.17	89	824
	1 MAS	E	E	E	E	E+LD+RP	148.28	9.73	45	432
Two-Stage	5 MAS	E+LD	E+RP	E+RP	E+RP	E+RP	14.83	7.81	166	2000
		E+LD	E+LD	E+RP	E+RP	E+RP	8.65	3.43	130	2000
		E+LD	E+LD	E+LD	E+RP	E+RP	7.86	1.68	94	2000
		E+LD	E+LD	E+LD	E+LD	E+RP	19.17	1.27	58	2000
	4 MAS	E	E+LD	E+RP	E+RP	E+RP	14.68	7.50	126	1608
		E	E+LD	E+LD	E+RP	E+RP	9.59	3.09	90	1608
		E	E+LD	E+LD	E+LD	E+RP	21.86	1.69	54	1608
	3 MAS	E	E	E+LD	E+RP	E+RP	16.38	7.54	86	1216
		E	E	E+LD	E+LD	E+RP	21.47	3.13	50	1216
	2 MAS	E	E	E	E+LD	E+RP	39.27	7.94	45	824
Combined	5 MAS	E+LD	E+LD+RP	E+LD+RP	E+LD+RP	E+LD+RP	21.16	1.33	182	2000
		E+LD	E+LD	E+LD+RP	E+LD+RP	E+LD+RP	30.99	1.26	142	2000
		E+LD	E+LD	E+LD	E+LD+RP	E+LD+RP	54.87	1.19	102	2000
		E+LD	E+LD	E+LD	E+LD	E+LD+RP	108.24	1.15	62	2000
	4 MAS	E	E+LD	E+LD+RP	E+LD+RP	E+LD+RP	40.53	1.61	138	1608
		E	E+LD	E+LD	E+LD+RP	E+LD+RP	64.04	1.49	98	1608
		E	E+LD	E+LD	E+LD	E+LD+RP	117.18	1.38	58	1608
	3 MAS	E	E	E+LD	E+LD+RP	E+LD+RP	69.41	2.25	94	1216
		E	E	E+LD	E+LD	E+LD+RP	117.05	1.94	54	1216
	2 MAS	E	E	E	E+LD	E+LD+RP	113.32	4.63	49	824

Refining the Strategy

To fine-tune the strategy, population size was considered as a possible means to reducing FR NRP. By increasing population size during the generations involving E+LD while continuing the same selection

intensity, the probability of recovering some individuals with less FR NRP is increased (Fig. 18). Recall that equilibrium for FR NRP is reached when residual NRP in the region is about 1 cM.

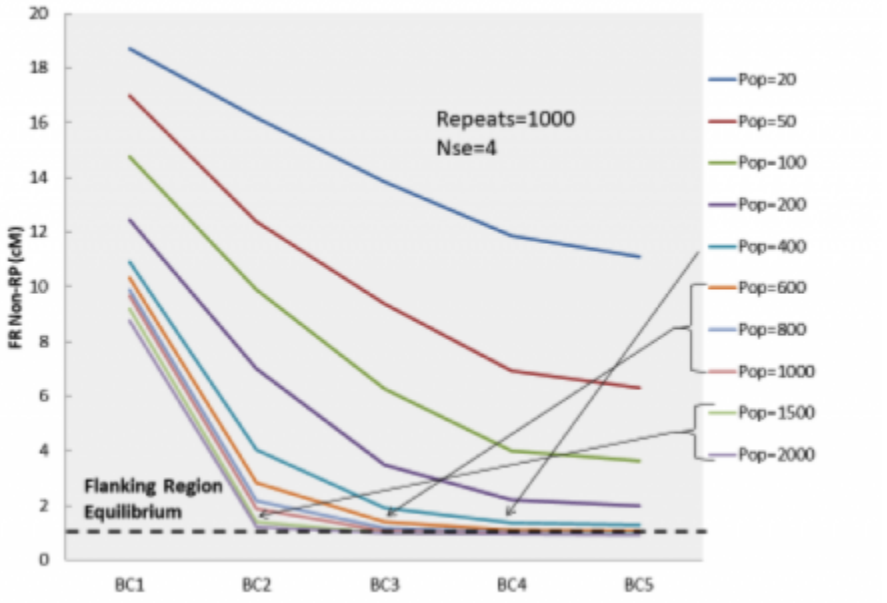


Fig. 18 Impact of population size on generation by which FR NRP equilibrium is reached. Adapted from Peng et al, 2014a.

Optimizing the Strategy

To build on the earlier simulation results, strategies with varied population sizes for the three generations of E+LD selection were considered. In conclusion, Peng et. al. (2014a) found that the strategy involving five generations of selection, the first three generations for E+LD with a population size of 600 followed by two generations of selection for E+RP with a population size of 400, resulted in 6.57 cM of Total NRP with 1.18 cM in the flanking region. Goal achieved!

This strategy required 100,600 marker data points and 2600 individual backcross progeny overall (Table 5). Costs for these could be estimated to provide a comprehensive view of the budget implications versus benefits of implementing this strategy in Trait Integration.

Table 5 Variations in population size to reach the goal of <8 cM Total NRP and ~1 cM FR NRP. Data from Peng et al., 2014a.

Strategy	BC_1 E+LD	BC_2 E+LD	BC_3 E+LD	BC_4 E+RP	BC_5 E+RP	Total NRP	FR NRP	MDP (K)	Ntotal
1	400	400	400	400	400	7.86	1.68	94.0	2000
2	600	600	600	400	400	6.57	1.18	100.6	2600
3	800	800	800	400	400	6.10	1.13	107.2	3200

Choosing Future Donor Parents

Of course, once "clean" conversions have been achieved, these can be used as trait donors in subsequent conversions. With only ~1 cM of non-elite DNA surrounding the event of interest (genetic factor), such conversions represent an excellent choice as donor parent to initiate new VAT conversions.

What other factors should be considered in choosing the donor parent for VAT conversions?

Genetic similarity to the line targeted for conversion would be beneficial since the goal of backcrossing is to recover the likeness of the RP. If the donor parent is related to the RP, then fewer backcross generations would be required to recover the RP attributes.

Computer simulation showed that if the NRP is 30% similar genetically to the RP, then complete conversion can be achieved by BC_4 using markers for selection of E, LD, and RP. If the NRP is 86% similar genetically to the RP, then complete conversion can be achieved by BC_3.

Refer to Molecular Plant Breeding for more details.

Next Steps

The results for Single Gene/Event Introgression indicated that the goal of a very "clean" conversion was achievable!

Furthermore, it sets the stage for optimizing the other steps in Multiple Trait Integration: Gene/event Pyramiding, Trait Fixation, and Version Testing (Fig. 17).

More on the Simulation Scenario

Continuing further with the breeding scenario to convert a corn hybrid for 15 events, some additional assumptions were added as next steps in Multiple Trait Integration were analyzed:

- Hybrid conversions involved stacking 8 events in the female parent $\mathbf{RP_F}$ of the hybrid and 7 events in the male parent $\mathbf{RP_M}$ of the hybrid.
- On each side of the pedigree, events are on different chromosomes (unlinked).
- Up to 5 versions of each converted hybrid were considered.
- The target hybrid yielded 14.72 tons per hectare (i.e. 235.6 bushels per acre) on average.
- The overarching objective is recovery of a converted hybrid that yields within 3% of the target hybrid.

The goals for each step are defined by the desired outcome. For example, the outcome of Single Event Introgression is *k* versions of each RP, each of the *k* versions representing a single event introgression into the RP background, created through backcrossing. With 8 events being stacked in the female parent of the hybrid, $k = 8$ for RP_F. At the close of Single Event Introgression, the backcross lines carry the desired event in a heterozygous state (Fig. 19).

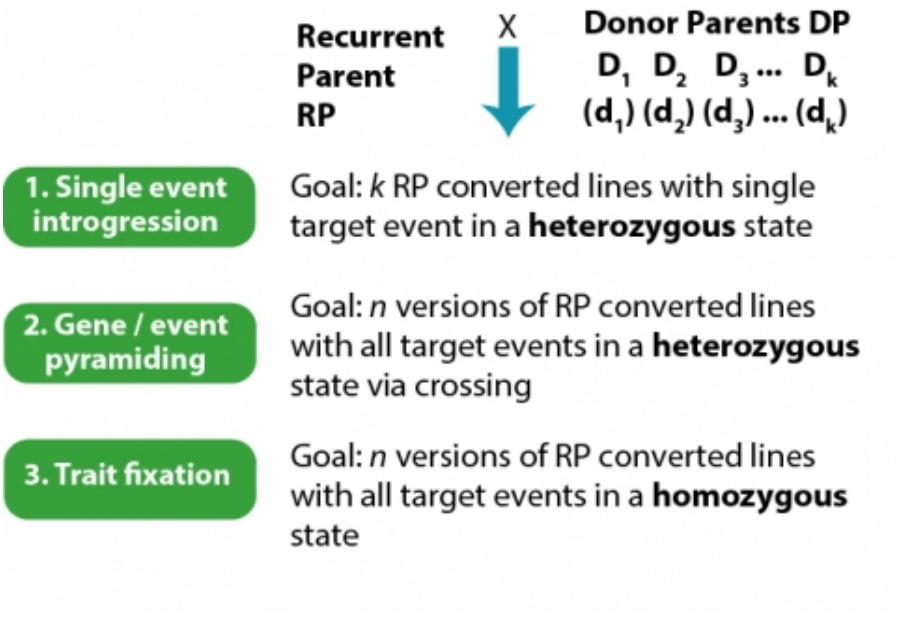


Fig. 19 Goals for Single Gene/Event Introgression, Gene/Event Pyramiding, and Trait Fixation. Adapted from Ishii and Yonezawa, 2007.

Event Pyramiding

Once quality single-event RP conversions have been created, the most efficient way to combine the events in the RP background is with a set of symmetrical crosses.

To stack 8 events in the RP_F: first, pairs of single event lines are crossed, then the double-event progeny are crossed, and finally the quad-event progeny are crossed to combine all 8 events in the RP_F background. Thus, 3 generations are required to combine all 8 events in heterozygous state in the specific RP.

Note that, although the simulation considered all events for a particular RP unlinked, in real life, linked events demand special attention in crossing to minimize the occurrence of repulsion linkages.

Trait Fixation

Theoretically, it would take a single generation of self-pollination to recover at least one progeny that is homozygous for all 8 events; each single event would be segregating in Mendelian fashion. However, the probability of recovering such an individual is very small.

Therefore, an extra generation may be introduced to alleviate bottlenecking due to extremely low frequency of the desired genotype. This concept is referred to as "**F_2 enrichment**" as per Bonnett et al. (2005). The outcome of the extra generation of self-pollination is it allows for individuals that have at least one copy of every event to be advanced (Fig. 20).

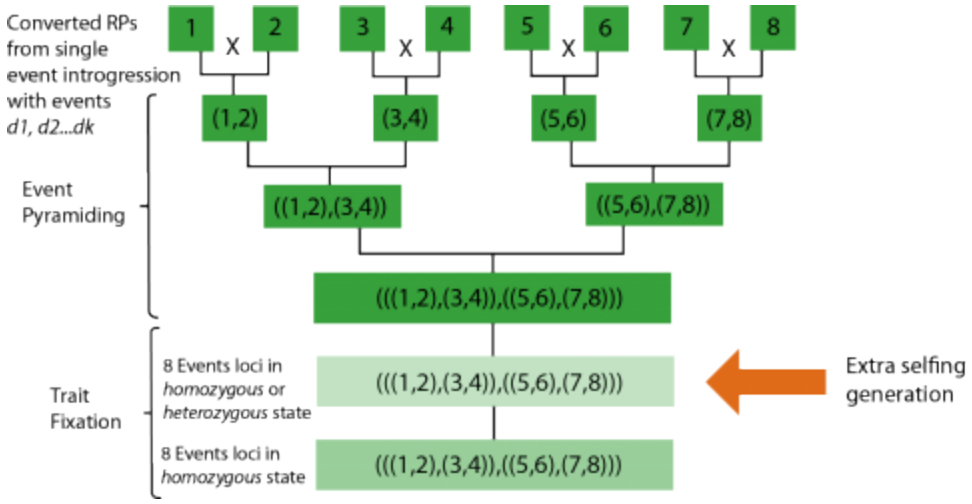


Fig. 20 Schematic showing pyramiding and fixing 8 events. Adapted from Peng et al, 2014b.

Practical Considerations

Two factors are very important to determining success in recovering specific desired genotypes: the population size and the expected frequency of the desired genotype in the population. Another factor is the number of individuals to be recovered. These factors must be stated for each generation.

The expected frequency of the desired genotype is a probability function based on Mendelian genetics. Note the example below (Table 6).

Table 6 The expected frequency of the desired genotype in each of the stacking and selfing generations to introgress 8 events. Peng et al., 2014b.

Generation	Pyramid 2 Events	Pyramid 4 Events	Pyramid 8 Events	S1 with 8 Events in heterozygous/ homozygous state	S_2 with 8 Events in homozygous state
Genotype/Event	Aa	Aa	Aa	AA/Aa	AA
Formula	$(0.5)^2$	$(0.5)^4$	$(0.5)^8$	$(0.75)^8$	$(0.5)^8$
Probability	0.25	0.0625	0.00390625	0.100112915	0.00390625

Determining Minimum Population Size

Knowing the expected frequency of the desired genotype and stipulating the number of individuals to be recovered as well as the probability of success, the necessary minimum population size can be determined.

Based on the binomial distribution (Sedcole 1977):

$$\sum_{i=x}^{N} \binom{N}{i} p^i (1-p)^{N-i} \geq q$$

where:
N refers to the minimal population size
x is the number of recovered individuals with the desired genotype
p is the probability of achieving the breeding goal
q is the frequency of the desired genotype in the population.

For the special case of *x*=1, the following formula applies:

$$N \geq ln(1-p)/ln(1-q)$$

To recover with 99% probability at least one plant with a genotype expected at the 0.25 frequency in the population,

$$N \geq ln(1-.99)/ln(1-.25)$$

$$N \geq 16.00784556$$

Thus, 99% of the time, 17 plants are needed to recover ≥1 with the desired genotype when that genotype is expected to occur in 1 in 4 individuals in the population.

Population Size Guide

Sedcole (1977) provided a handy table to help in quickly estimating the population size needed to guarantee recovery of desired genotypes with high probability (Table 7).

Table 7 Number of plants necessary to recover a required number of plants with trait (r = number of plants to be recovered).

p*	q^	1	2	3	4	5	6	8	10	15
0.95	1/2	5	8	11	13	16	18	23	28	40
1/3	8	13	17	21	25	29	37	44	62	n/a
1/4	11	18	23	29	34	40	50	60	84	n/a
1/8	23	37	49	60	71	82	103	123	172	n/a
1/16	47	75	99	122	144	166	208	248	347	n/a
1/32	95	150	200	246	291	334	418	500	697	n/a
1/64	191	302	401	494	584	671	839	1002	1397	n/a
0.99	1/2	7	11	14	17	19	22	27	32	45
1/3	12	17	22	27	31	35	44	52	71	n/a
1/4	17	24	31	37	43	49	60	70	96	n/a
1/8	35	51	64	77	89	101	124	146	198	n/a
1/16	72	104	132	158	182	206	252	296	402	n/a
1/32	146	210	266	218	268	416	508	597	809	n/a
1/64	293	423	535	640	739	835	1020	1198	1623	n/a

* = probability of recovering r plants with trait.
^ = probability of occurrence of trait.

Further Efficiencies in Trait Fixation

Further efficiency in Trait Fixation can be considered in terms of plant materials for genotyping (e.g. seed chipping before planting or sampling leaf tissue of seedling plants). Bottom line: with seed chipping, selection can be performed *before planting* so resources are focused only on the individuals with the desired genotype.

In addition, cycle time is decreased because the S_2 outcome of selection is determined before planting. If testcrossing is required to facilitate version testing, the S_2 materials can be planted appropriately to produce the hybrids.

Trait Fixation also results in the production of Breeder Seed to hand-off to Supply Chain. Those individuals deemed to be fixed for the introgressed genetic factors will be self-pollinated to produce seed which can be bulked to comprise the converted new elite line. The hand-off to Supply Chain will occur once the performance equivalency of the converted line has been confirmed.

Further Efficiencies Overall Through Rapid Cycling

In addition to accomplishing the breeding goal each generation, the entire process of Trait Integration can be

accelerated through rapid cycling of the generations for backcrossing, pyramiding, and trait fixation. In other words, ways to achieve more than one generation per year speed the process.

This can be accomplished through the use of off-season nurseries, continuous nurseries in tropical locations, or the greenhouse to facilitate the continual rapid cycling of generations. With corn, continuous nurseries have enabled the cycling of four generations per year. Methods of intensive nursery management resulting in a rapid progression of generations have been referred to as **speed breeding**.

Version Testing

For the final step in Multiple Trait Integration, the converted cultivar is tested to assess its performance against its unconverted counterpart. The goal is performance equivalency.

The factors that determine success are:

- The amount of residual NRP germplasm in the converted cultivar.
- The probability of successful recovery of at least one version of the cultivar with equivalent performance.
- The number of stacked versions of each RP.

Just because the conversions have been created, it does not mean that performance equivalency has been recaptured. However, Sun and Mumm (2015) showed that, given the strict standards of single event conversions shown, there is a high likelihood of recovering an equivalent conversion even when the conversion involves 15 stacked events. Encouraging!

See Sun and Mumm (2015) for more details.

What's Next?

This chapter has focused on the Trait Integration component of the product pipeline, an important element in cultivar development. In addition, it has featured results of a study in process design for Trait Integration, because efficiency that leads to a higher rate of genetic gain is a critical part of getting better cultivars into farmers' hands.

The next chapter will consider ways to further optimize the product pipeline.

References

Bonnett, D.G., G.J. Rebetzke, W. Spielmeyer. 2005. Strategies for efficient implementation of molecular markers in wheat breeding. Molecular Breeding. 15:75-85.

Comai, L., S. Henikoff. 2006. TILLING: Practical single-nucleotide mutation discovery. Plant J. 45:684-694. DOI: 10.1111/j.1365-313X.2006.02670.x

Fehr, W. R. 1987. Principles of Cultivar Development Volume 1: Theory and Technique. McMillan Publishing Company.

Ishii, T., and K. Yonezawa. 2007a. Optimization of the Marker-Based Procedures for Pyramiding Genes from Multiple Donor Lines: I. Schedule of crossing between the donor lines. Crop Sci. 47(2):537-546.

Ishii, T., and K. Yonezawa. 2007news.psu.edu/.../gene-edited-mushroom-created-penn-state-researcher-changing-gmob. Optimization of the Marker-Based Procedures for Pyramiding Genes from Multiple Donor Lines: II. Strategies for selecting the objective homozygous plant. Crop Sci. 47:1878-1886.

James, Clive. 2015. 20th Anniversary (1996 to 2015) of the Global Commercialization of Biotech Crops and Biotech Crop Highlights in 2015. ISAAA Brief No. 51. ISAAA: Ithaca, NY.

Monsanto-DowAgrosciences Collaborative Agreement Presentation, September 14, 2014.

Mumm, R.H. 2013. A look at seed product development with genetically modified crops: Examples from maize. J. Agri. and Food Chem. 61(35):8254-8259. DOI: 10.1021/ jf400685y

Peng, T., X. Sun, and R.H. Mumm. 2014a. Optimized breeding strategies for multiple trait integration: I. Minimizing linkage drag in single event introgression. Molecular Breeding 33:89-104. DOI 10.1007/ s11032-013-9936-7

Peng, T., X. Sun, and R.H. Mumm. 2014b. Optimized breeding strategies for multiple trait integration: II. Process efficiency in event pyramiding and trait fixation. Molecular Breeding 33:105-115. DOI 10.1007/ s11032-013-9937-6

Sun, X., T. Peng, and R.H. Mumm. 2011. The role and basics of computer simulation in support of critical decisions in plant breeding. Molecular Breeding 28 (4):421-436. http://www.springerlink.com/openurl.asp?genre=article&id= doi:10.1007/s11032-011-9630-6

Sun, X., and R.H. Mumm. 2015. Optimized breeding strategies for multiple trait integration: III. Parameters for success in version testing. Molecular Breeding 35:201. DOI 10.1007/s11032-015-0397-z

Till, B.J., S.H. Reynolds, C. Weil, N. Springer, C. Burtner, K. Young, E. Bowers, C.A. Codomo, L.C. Enns, A. R. Odden, E. A. Greene, L. Comai, S. Henikoff. 2004. Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biology 4:12. DOI:10.1186/1471-2229-4-12

How to cite this chapter: Mumm, R.H. (2023). Value-added Trait Integration. In W. P. Suza, & K. R. Lamkey (Eds.), *Cultivar Development*. Iowa State University Digital Press.

Chapter 6: Optimization of Product Pipeline

Rita H. Mumm

Response to Selection, *R*

The *cycle* of cultivar improvement is implemented as a *process* of cultivar development, which is designed to produce a *pipeline* of improved cultivars that achieve the specific product target (Fig. 1).

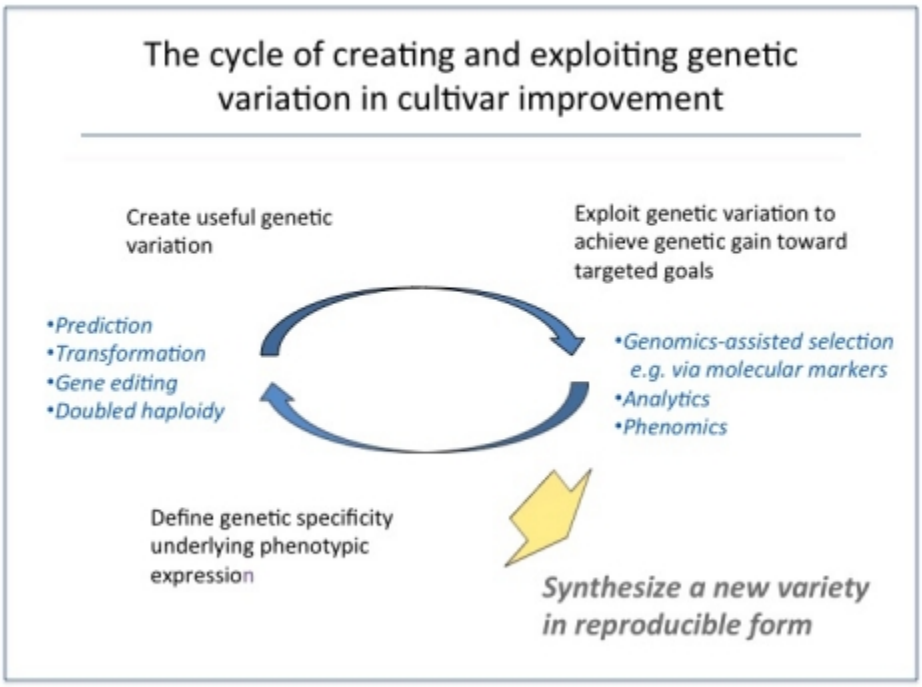


Fig. 1 Inclusion of modern tools in the cycle of creating and exploiting genetic variation in cultivar improvement. Adapted from Mumm, 2013,

Generally speaking, top-performing individuals are selected from a base population with mean performance, μ, to serve as parents. Progeny produced from these parents are evaluated, with the best selected to serve as parents of the next cycle of selection.

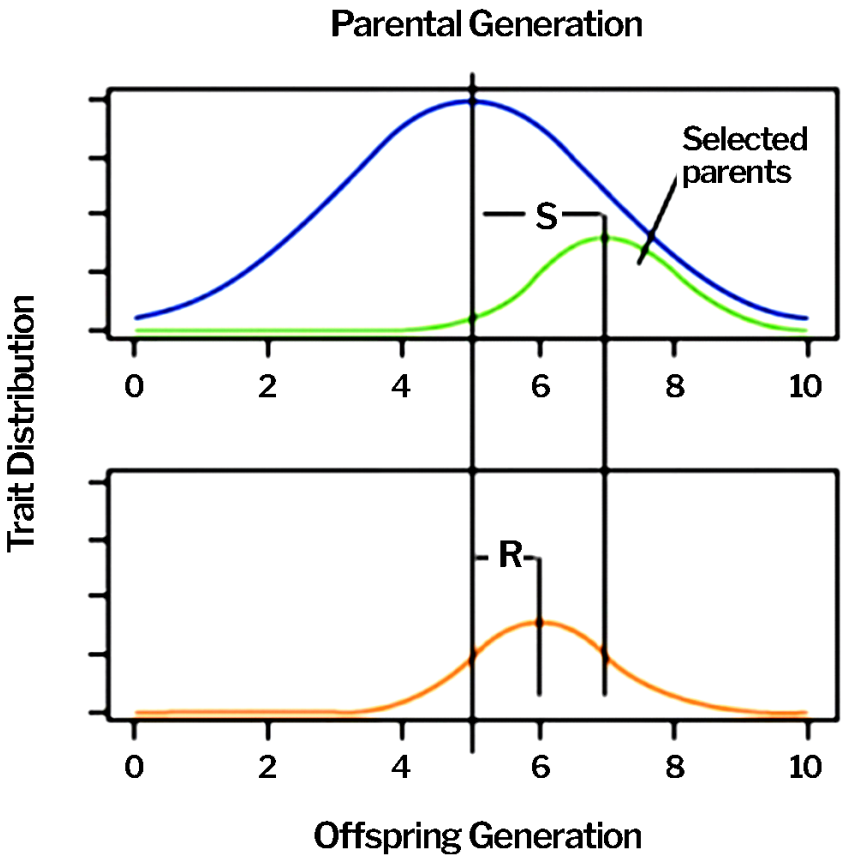


Fig. 2 Shifts in trait means from base population in parental generation and progeny generation as a result of selection.

Narrow sense heritability (h^2) is the portion of the genetic variance that can be transmitted to the next generation. It reflects the ratio of ***R*** to ***S***:

$$h^2 = \frac{R}{S}$$

Factors Influencing R

The efficiency of the process can be measured in terms of *R*, which represents the response to selection.

$$R = h^2 S$$

We have seen in earlier chapters that *R* can also be expressed as:

$$R = ih^2\sigma_P$$

$$R = i\left(\frac{\sigma_A^2}{\sigma_P^2}\right)\sigma_P$$

$$R = \frac{i\,\sigma_A^2}{\sigma_P}$$

$$R = ih\sigma_A$$

where:

i is selection intensity,

h^2 is narrow sense heritability for the trait(s) under selection,

σ_A^2 is the additive genetic variance,

σ_P^2 is the phenotypic variance,

A is additive genetic standard deviation,

P is phenotypic standard deviation,

h is the square root of h^2 and refers to the accuracy of selection.

Furthermore, the rate of genetic gain (**ΔG**) over time can be expressed as:

$$\Delta G = \frac{R}{L}$$

where L is the length of a breeding cycle.

Which of the factors influencing ***R*** and ***ΔG*** are impacted by choices the breeder makes?

R as an Efficiency Indicator

Optimization of the product pipeline promotes maximal response to selection in the shortest amount of time at a comparative cost to produce improved cultivars that meet the needs of farmers and end-users.

Through process design, the breeder can affect the efficiency of the pipeline in developing improved cultivars. *R* is an efficiency criterion that can be used to compare various process design options to create a pipeline that maximizes Δ*G* over time.

Let's take a look at how to utilize *R* as a measure of process efficiency…

More on Factors Influencing R

We have seen *R* expressed in a number of ways and these can be expanded further to elucidate more features of the breeding and testing regime over which the breeder has control.

For example:

$$R = h^2 S$$

where S is the selection differential, $S = \mu_S - \mu$

$$R = \frac{\sigma_A^2}{\sigma_P^2} S$$

$R = \frac{\sigma_A^2}{\sigma_P^2} k\sigma_P$ where k is the number of phenotypic standard deviations between $\mu_S - \mu$

$$R = \frac{k(c)(f)\sigma_A^2}{\sigma_P}$$

where:
c = parental control factor
f = fraction of σ_A^2 among progeny being tested.

e.g., $HS = \frac{1+F}{4}$ where F is the inbreeding coefficient of half sibs.

Impact of Parental Control

The parental control factor (*c*) reflects the relationship between the entity used for identifying genotypes (selection unit) and the entity used to produce the next generation (recombination unit).

Note that:

- ***c* = ½** when the selection unit is the same as the recombination unit and only female parents are selected
- ***c* = 1** when the selection unit is the same as the recombination unit and both parents are selected
- ***c* = 2** when the selection unit is not the same as the recombination unit; the recombination unit is the selfed seed or a clone of the selected.

For example, mass selection practiced with pollen control is twice as effective as without pollen control, all other factors being equal:

$$R = \frac{k(1)(1)\sigma_A^2}{\sigma_P} \text{ (with pollen control) vs. } R = \frac{k(\frac{1}{2})(1)\sigma_A^2}{\sigma_P} \text{ (without pollen control)}$$

It is easy to see that by conducting controlled pollinations using only selected lines, twice the genetic gain

can be made in comparison to recombining female parents that have produced seed through uncontrolled pollination. This can be done when selections (for both male and female lines) are made prior to pollination.

Impact of the Form of the Selection Unit

In the expression of R, the variable f has to do with the form of the selection unit. If half-sib progeny produced from the breeding crosses are subject to evaluation, then the proportion of the additive genetic variance represented is ¼. Likewise, if full-sib progeny are evaluated, then ***f* = ½**.

For example, full-sib selection is more effective than selection among half-sib progeny, all other factors being equal:

$$R = \frac{k(1)(\frac{1}{2})\sigma_A^2}{\sigma_P} \text{ (full-sib selection) vs. } R = \frac{k(1)(\frac{1}{4})\sigma_A^2}{\sigma_P} \text{ (half-sib selection)}$$

Full-sibs account for twice the additive variance compared to half-sibs.

Furthermore, f considers the level of inbreeding (F) among the progeny. F is assumed to be zero in the F_2 generation; whereas, with fully homozygous progeny, $F = 1$.

Assuming epistasis to be negligible, the genetic variability among families with or without inbreeding depends on the types of families under selection.

- **Half-sib** = $\frac{1+F}{2}\sigma_A^2$.
- **Full-sib** = $\frac{1+F}{2}\sigma_A^2 + \frac{(1+F)^2}{4}\sigma_D^2$.
- **Selfed** = $(1+F)\,\sigma_A^2 + \frac{1}{4}(1-F)(1+F)\,\sigma_D^2$; (Note that, with fully homozygous inbreds such as doubled haploid lines, this expression reduces to $2\sigma_A^2$).
- **Clones** = $\sigma_A^2 + \sigma_D^2 + \sigma_I^2$.

The numerator of R considers only the additive variance because only additive variation can be transmitted to progeny in a diploid species. However, other forms of genetic variation (e.g. dominance variance) contribute to the denominator of R which we will explore later in this chapter.

The Impact of k on R

With k denoting the number of phenotypic standard deviations that the mean of the selected differs from the base population mean, k is the same as i (selection intensity; shown in earlier expressions of R). Although individuals in the extremes of the distribution may be superior, they are rare. The chance of producing,

identifying, and selecting individuals in the extremes of the distribution curve is enhanced as the number of lines tested increases. It is difficult to increase k without concomitantly increasing the population size. Otherwise, the risk is depleting genetic diversity.

This relationship is illustrated in the following example: 10% selection from a population size of 10 means that 1 individual is selected; no genetic diversity remains for the next round of selection.

In contrast, 10% selection from a population size of 250 means that 25 individuals are selected. Additionally, a sufficient number of selected individuals to be recombined for the next cycle of selection contributes to reducing the potential effects of genetic drift.

In fact, k is related to p, the proportion of selected individuals in the base population. Falconer (1989; Appendix Table A) shows the relationship between i (i.e. k) and p and x, the latter being the difference between the threshold point of selection (i.e. truncation point) and the base population mean, expressed in standard units.

Plant breeders often set p at certain values; p translates to values of x in the Z Table. Note that x for 5% and for 1% are the familiar statistics from the Z Table, 1.645 and 2.326, respectively (Table 1).

Table 1 Some commonly used values of p, along with associated values of x and i. From Falconer, 1989.

p (%)	x	i (i.e. k)
0.1	3.090	3.367
0.5	2.576	2.892
1	2.326	2.665
2	2.054	2.421
5	1.645	2.063
7.5	1.440	1.887
10	1.282	1.755
25	0.674	1.271
50	0	0.798

Other Ways to Maximize R

So far we have looked at ways to maximize R by enlarging the numerator in the R equation:

$$R = \frac{k(c)(f)\,\sigma_A^2}{\sigma_P}$$

In addition, R can be optimized by reducing the denominator.

Partitioning Phenotypic Variation

We know that phenotypic variation is a function of variation attributable to genotype, environment, genotype-by-environment interaction (*GxE*), and error.

Simply put, the denominatorσ_P can be expressed as:

$$\sigma_P = \sqrt{\frac{\sigma^2}{rt} + \frac{\sigma^2_{GE}}{t} + \sigma^2_G}$$

where:
σ^2 is experimental error
σ^2_{GE} is variance due to genotype by environment interaction
σ^2_G is genetic variance
r is the number of replications in testing
t is the number of environments (which could include locations, seasons, years, cultural practices, etc.).

Accordingly, this expression can be expanded, with σ^2_P expressed as:

$$\sigma_p = \sqrt{\frac{\sigma^2_w}{nrly} + \frac{\sigma^2_e}{rly} + \frac{\sigma^2_{GLY}}{ly} + \frac{\sigma^2_{GY}}{y} + \frac{\sigma^2_{GL}}{l} + \sigma^2_G}$$

where:
σ^2_w is the within-plot experimental error variance,
σ^2_e is the variance among replications,
σ^2_{GLY} is variance due to genotype by location by year interaction,
σ^2_{GY} is variance due to genotype by year interaction,
σ^2_{GL} is variance due to genotype by location interaction,
σ^2_G is genetic variance, $n$is the number of plants per plot,
r is the number of replications in testing,
l is the number of testing locations
y is the number of years of testing.

Note that the mean square error in ANOVA is comprised of within-plot and among-plot variation:

$$\frac{\sigma^2}{rt} = \frac{\sigma^2_w}{nrly} + \frac{\sigma^2_e}{rly}$$

Moreover, the within-plot experimental error variance σ_w^2 can be partitioned further into variation due to environmental effects (σ_u^2) and variation due to genetic differences among plants (σ_{wg}^2):

$$\sigma_w^2 = \sigma_u^2 + \sigma_{wg}^2$$

Environmental effects σ_u^2 include micro-scale effects within the plot that would cause genetically identical plants to perform differently (e.g. soil fertility, soil type, soil moisture, shading). Within-plot genetic variation σ_{wg}^2 is attributable to segregation such as might occur before lines are fully inbred.

In addition, variation attributable to GxE can be expanded. Thus,

$$\frac{\sigma_{GE}^2}{t} = \frac{\sigma_{GLY}^2}{ly} + \frac{\sigma_{GY}^2}{y} + \frac{\sigma_{GL}^2}{l}$$

The variation attributable to GxE could be expanded further if needed to include variance due to 'seasons' or cultural practices (e.g. irrigation vs. dryland).

Maximizing R Through Trial Design

R can be maximized through design of evaluation trials. Uniform fields will contribute to less variation within plots and among replications, thus smaller σ_w^2 and σ_e^2, respectively.

To maximize R, σ_P can be reduced by increasing the number of reps (r), number of locations (l), and/or number of years (y) in evaluation for selection. Relatively speaking, increasing locations or years in testing has a greater effect than increasing replications (that is, l and y have the opportunity to reduce more components of variation since they are featured in the denominators more often than r).

Rwill increase as the number of plants per plot (n) is increased. When there is only 1 plant per plot, $n = 1$. In family selection, the value of n is determined by the number of plants in the plot. For example, with F_3 family selection involving 30 plants per plot, $n = 30$. The effect of increasing plant numbers per plot as it relates to $\frac{\sqrt{\sigma_w^2}}{n}$ can be seen in the following table (Table 2).

Furthermore, error variance components, σ_w^2 and σ_e^2, can be reduced by controlling human error such as mistakes in recording the evaluation data.

Table 2 The effect of increasing n with constant σ_w^2.

n	$\frac{\sqrt{\sigma_w^2}}{n}$
1	22.3
2	15.8
3	12.9
4	11.2
5	10.0
10	7.1
20	5.0
30	4.1
60	2.9
90	2.4

The Impact of GxE on R

The impact of GxE can be reduced by evaluating progeny in multiple locations and over multiple years. Variation due to GxE is captured in σ_{GLY}^2 (variance due to genotype by location by year interaction), and σ_{GY}^2 (variance due to genotype by year interaction), and σ_{GL}^2 (variance due to genotype by location interaction). Because GxE interferes with the ranking of the progeny and identification of top performers, it has serious ramifications for accuracy in selection.

For crops in which GxE features largely, the greatest genetic gain per year is realized by maximizing l at the expense of r (i.e. 1-rep trials at the greatest number of locations possible). The trade-off here is cost, since with a fixed number of reps, it is more expensive to include more locations than to have multiple reps at fewer locations. (More on choice of locations later in this chapter.)

Note that σ_{GY}^2 and σ_{GLY}^2 cannot be measured or effectively reduced without testing over multiple years. Obviously, the trade-off is time. As an alternative option, additional locations may be substituted for added years.

Choosing Testing Environments to Maximize R

Choice of testing sites is critical to the selection decisions that will be made in the process of cultivar development. Selections will be made based on phenotypic performance at these sites for traits pertinent to the product target.

Prospective testing sites are assumed to represent a sample of environments from the target market

region. However, realistically there may be hundreds of thousands of “environments” within the target region, considering the many factors that play a role in determining the environment (e.g. geography, altitude, season, maturity zone, soil types, topography, farmer production management system including tillage, fertilization, mechanization, and water regime, etc.). Selection in cultivar development is aimed at the identification of genotypes that have high mean value in the target set of environments.

The usefulness of a testing site can be characterized in terms of:

- Ability to differentiate between genotypes
- Accuracy of selection: h
- Correlation of test environment with target environment within the market region.

The purpose of testing in cultivar development is to reveal genetic differences among test entries and the magnitude of those differences, and to indicate whether performance of any is sufficient to meet the stated product target. Trial design and implementation have an important role in this. However, some types of environments (e.g. drought stress and other abiotic stress environments) inherently have greater variability due to non-genetic factors.

Heritability (h^2) can be considered a measure of the signal-to-noise ratio in estimating breeding value; GxE interaction and experimental error reduce the effectiveness of selection:

$$h^2 = \frac{\sigma_A^2}{\frac{\sigma^2}{rt} + \frac{\sigma_{GE}^2}{t} + \sigma_G^2}$$

Accuracy of selection or h is defined as the correlation between breeding value and phenotypic value and can be computed as the square root of narrow-sense heritability ($\sqrt{h^2}$) (Falconer 1989); it is considered a measure of repeatability. Therefore, the accuracy of selection decisions depends on reducing GxE and experimental error in testing. Accuracy of selection often differs across testing sites, but this statistic alone is not sufficient to establish the value of a testing environment.

Because selections are based on a sample of the entire population of environments, selection will be truly accurate only if the observed phenotype in that particular environment or set of environments is highly correlated with mean performance in the target environment.

Allen et al. (1978) suggested that an appropriate measure of the usefulness of a test environment is $r\sqrt{h^2}$ where r is the genetic correlation of performance in the test environment with performance in the target environment.

Challenges in Choosing Testing Environments

Ideally, test environments are discriminating among genotypes as well as representative of the target market region.

Some questions to consider in identifying useful test sites are:

- Are there multiple target environments (i.e. **mega-environments**) within the target market region?
- Within a specific mega-environment, what are the most representative and discriminating test sites?
- How many test sites (and replications within a trial) can be implemented each year on my budget?

With a defined budget, it is important to understand meaningful mega-environments within the market region and to utilize testing environments that provide the most useful information in identifying "best" genotypes.

GxE Interaction Patterns

Ideally, an improved cultivar with wide adaptation across the entire defined target market is developed. This is relatively easy if GxE is negligible across the market region. In this case, test entries would perform similarly across test sites; differences between entries would not change from one environment to the next, and thus rankings would not change.

Sometimes the level of GxE is such that differences between lines change in magnitude from one environment to the next, yet rankings of entries based on performance still hold.

Crossover interaction is evident when differences between entries change substantially, resulting in rankings being flipped. A top entry at one location may be the worst-performing entry at another (Fig. 2).

- Parallel pattern: no interaction.
- Non-parallel but rankings still hold.
- Crossover interaction: Rankings are flipped!

With consistent patterns of GxE over years, there are groups of locations that consistently share the best set of genotypes or cultivars, and this implies a repeatable pattern that can be used to exploit GxE in cultivar development. Strategically, it means that selection will focus on specifically adapted genotypes for each mega-environment. Thus, each mega-environment becomes a "target environment" and test sites within that mega-environment must be representative of that target environment.

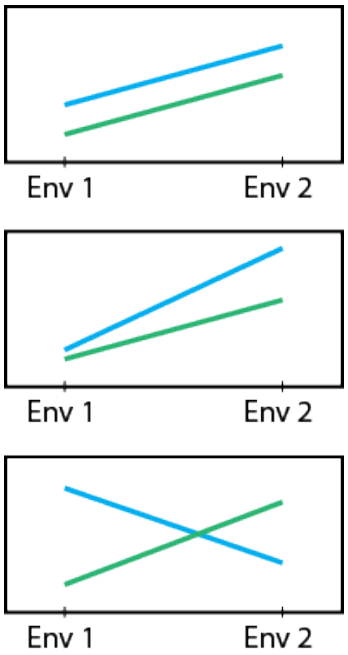


Fig. 2 Patterns of GxE: parallel, non-parallel, and crossover interaction.

Tools to Aid in Evaluating Testing Environments

There are a number of tools that can aid in answering the questions posed (see slide entitled Challenges in Choosing Testing Environments in this chapter).

Cluster analysis has been used to identify similar types of environments and categories of environments. In addition, the performance of test entries grown at one location can be correlated with the performance of the same set of entries at other locations and performance overall.

AMMI (Additive Main Effects and Multiplicative Interaction) analysis has been used to dissect GxE via principal components analysis and characterize yield stability over environments (Gauch 1992). A biplot can be produced to graphically display the relationships of environments with each other and with test entries.

GGE biplot involves only genotype and GxE interaction effects in the principal components analysis. A biplot displaying environments and entries based on the first two principal components, PCA1 and PCA2, provide insights into the number of mega-environments represented in the data and indicates test environments that are most representative and discriminating. Graphics can be unscaled or scaled in different ways (e.g. by the standard deviation of an entry mean within environment or the standard error within environments).

In particular, when GGE biplots are scaled according to the heritability of the trait measured, the length of the vector extending from the origin to the environment is proportional to the square root of heritability ($r\sqrt{h^2}$) in the environment. This type of analysis is referred to as a heritability-adjusted GGE biplot (Yan and Holland 2010) (Fig. 3). The angles between environment vectors approximate the genotypic correlation between environments.

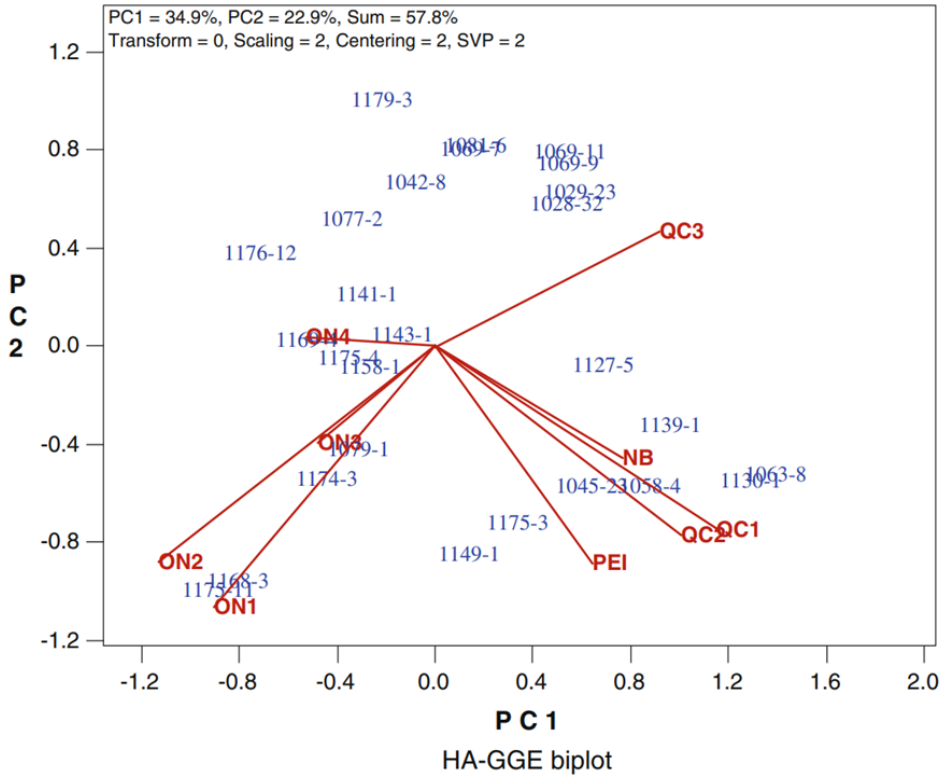


Fig. 3 Heritability-adjusted GGE biplot facilitates visualization of \(\sqrt{h^2}\) for 9 testing environments and the correlations among them (environment vectors in red). From Yan and Holland, 2010.

Furthermore, this type of analysis provides approximations of $r\sqrt{h^2}$ for each test environment, which is proportional to the predicted R in the mega-environment based on the data from the test environment. Thus, it delivers information the breeder can use in choosing environments that maximize R in a way that fits within budgetary constraints.

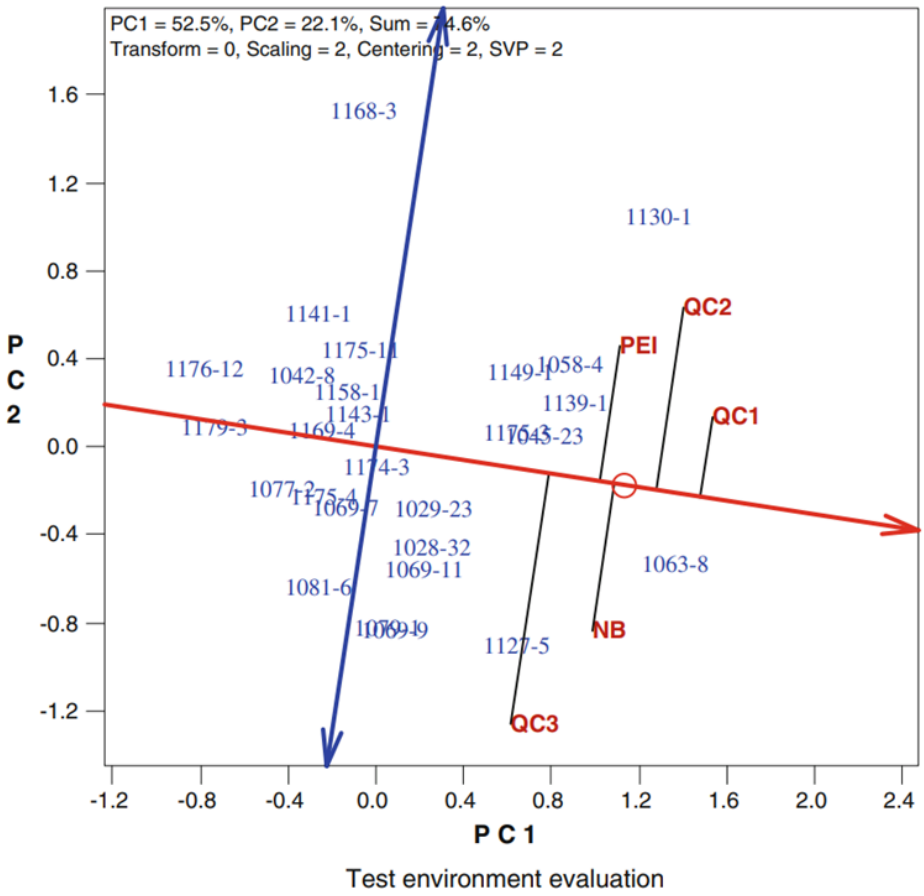


Fig. 4 Heritability-adjusted GGE biplot displays the usefulness and representativeness of the five environments associated with one mega-environment. Usefulness ranking given as QC1>QC2>NB>PEI>QC3. Environments with shorter lines to Target Environment Axis (TEA) are more representative; QC3 was least representative. From Yan and Holland, 2010.

Impact of Type of Gene Action on R

In the expression of R, the denominator includes all the genetic variance while only the additive genetic variance is included in the numerator (since only additive genetic variance can be transmitted to progeny):

$$R = \frac{k(c)(f)\sigma_A^2}{\sqrt{\frac{\sigma_w^2}{nrly} + \frac{\sigma_e^2}{rly} + \frac{\sigma_{GLY}^2}{ly} + \frac{\sigma_{GY}^2}{y} + \frac{\sigma_{GL}^2}{l} + \sigma_G^2}}$$

The genetic variation accounted for in the denominator will depend in part on the gene action of the trait being improved. This means that traits that are substantially influenced by dominant or epistatic gene action are generally less responsive to phenotypic selection. And generally, such traits are characterized by lower heritability. Typically, epistasis is assumed to be negligible in estimating R.

In addition, the family structure of the population under improvement will impact R. As shown above, the genetic variability among families with or without inbreeding depends on the types of families under selection. For example, half-sib families reflect additive genetic variance only, whereas full-sib families reflect both additive and dominance genetic variance. In estimating R using half-sib families, only additive genetic variance would be included in the denominator.

See Fehr, 1991, Chapter 17 for examples of comparing alternative breeding methods for predicted gain (R).

Breeders cannot change gene action. However, there are ways to deal with low heritable traits, including the use of indirect selection through secondary traits that may offer higher selection efficiency.

Use of Indirect Selection to Increase R

This really relates to the type of screen that is utilized to select for a given trait in the product target. Generally speaking, the screen should be accurate and relevant to field performance. As discussed in Chapter 2, to improve Character X, selection may be based on another Character Y. Indirect selection will result in a relatively greater genetic gain for Character X than directly selecting for it if:

$$r_{G_{XY}} h_Y > h_X$$

where:
$r_{G_{XY}}$ is the genetic correlation between Characters X and Y,
h_X and h_Y represent the square root of the narrow sense heritability for Characters X and Y, respectively.

Marker-assisted selection (MAS) is one type of indirect selection. Selection for superior individuals is based on the genotype of specific marker(s) instead of phenotype (or in addition to phenotype). Since marker genotypes are completely heritable (thus, h_Y is 1), the efficiency of marker-based selection is a function of the heritability of the primary trait and the proportion of additive genetic variance of the primary trait attributable to the markers.

When the proportion of additive genetic variance accounted for by the markers is greater than the heritability of the primary trait, then $r_{G_{XY}}$ is greater than h_X, and MAS holds an advantage.

Impact of L on ΔG

Response to selection over time is given by the equation for the rate of genetic gain:

$$\Delta G = \frac{R}{L}$$

Obviously, the length of the breeding cycle , **L**, is a significant element in maximizing **G** .

What are ways in which **L** can be reduced?

1. By increasing the number of generations advanced per year in breeding to homozygosity. This may involve use of growth chambers, greenhouses, **off-season nurseries**, and **continuous nurseries** that facilitate more than one generation per year and may involve breeding methods like single-seed descent.
2. Use of trait screens that facilitate selection before pollination (so that both parents are from the selected group) or even facilitate selection before planting (marker-based selection using plant seed tissue for DNA extraction).
3. Reducing the length of the time required for the testing regime. This may involve a trade-off between the number of years of testing and the number of locations grown each year in order to maintain a comparable R.

These ways are but some of the many ways that L can be reduced.

Can you think of others?

Check out this YouTube video that shows approaches to speed breeding devised by the Hickey Lab at the University of Queensland.

Accelerating Development Time

Just as development time can be reduced by shortening the length of a generation, still another way to realize efficiency is in terms of time to develop an improved cultivar that reflects a specific magnitude of genetic gain. After all, we are interested in releasing improved cultivars to farmers in the shortest amount of time required, depicted in the equation below.

$$Time = R * L$$

Development time required involves a trade-off between the magnitude of R per cycle (L) and the number of cycles of selection. If greater gain can be achieved per cycle, fewer cycles (and less time) may be needed to achieve a particular genetic gain.

Having discussed factors to increase R, one that is recognized for reducing time to market is population size. With enough increase in population size, the gain achievable in two cycles of selection could potentially be realized in one.

Recap

Study of the formulas for R has revealed several ways for increasing response to selection:

1. Select a smaller proportion of individuals; coordinate higher selection intensity with increased population size to avoid loss of genetic diversity and potential effects of genetic drift.
2. Utilize breeding methods that allow for control of both parents; choice of pollen parent as well as seed parent to produce progeny for the next cycle of selection offers great advantage. Furthermore, utilizing approaches to allow recombination of selfed seed or clones of selected individuals (vs. remnant seed)

adds to greater progress.

3. Increase the coefficient of σ_A^2; f can be increased through the choice of selection unit (e.g., full-sib families vs. half-sib families) and by evaluating more highly inbred materials (e.g. recombinant inbred lines vs. early generation segregating lines).
4. Deploy appropriate experimental designs in testing to minimize environmental variation and error. For advanced yield trials involving high numbers of entries, incomplete block designs can be used to further partition blocks in the field into more homogeneous units. Include multiple locations and multiple years to partition and account for variation due to GxE. Increase replications to the greatest extent feasible.
5. Control within-plot variation through appropriate plot size (i.e. number of plants per plot). Conduct single-plant evaluations only with highly heritable traits.
6. For lowly-heritable traits, consider the use of secondary traits as a basis for selection. Molecular markers may be useful if the markers aid in accounting for a greater portion of the additive genetic variance than phenotypic selection alone.
7. Reduce the length of the cycle by shortening the time required for the breeding and testing of progeny.
8. Reduce the number of cycles needed to achieve a particular magnitude of gain by increasing selection intensity through increased population size.

Thus, the breeder exerts tremendous influence on the response to selection and the rate of gain, simply through the choice of breeding strategies, methods, materials for testing, and the conduct of evaluations. Therefore, the design of the product pipeline which incorporates these features in the breeding and testing regime is critical to developing and releasing improved cultivars that meet farmer and end-user needs.

The breeder can utilize the formulas for R to compare various options in the product pipeline for the greatest efficiency.

Improving Breeding Efficiency Through the Use of Technology

The term “technology” refers loosely to industrial science. That is, any applied science in the form of tools, machines, or methodologies used to solve real-world problems. In crop improvement, technology includes DNA-based approaches to selection as well as the tools and methods to accomplish the selection, machines such as drones to record data more efficiently, and processes such as doubled haploidy that are enabled through scientific application.

Two technologies are highlighted here as examples, considering their potential role in the product pipeline and ways to maximize the benefit of the technological investment:

- **molecular markers**, specifically applied to facilitate indirect selection
- **doubled haploidy** to accelerate development of homozygous lines from a breeding cross.

Boosting Breeding Efficiency Through Marker-Assisted Selection

Molecular markers can be used to select for individuals with favorable alleles at loci controlling traits of interest in close proximity to marker loci. Selection based on molecular marker genotype may contribute to breeding efficiency.

Fig. 5 Photo by Scott Bauer. Courtesy of USDA ARS.

In potato, for example, Marker-Assisted Selection (MAS) can be used to address some of the critical challenges in cultivar development. Potato is a highly heterozygous, tetraploid crop with very complex product targets involving as many as 40 characteristics including disease and pest resistances, tuber appearance, quality, nutrient content, and stress tolerances, in addition to agronomic traits of yield, yield stability, and maturity (see Chapter 4). Evaluation is costly and time consuming. The process of New Line Development and New Line Evaluation can take 10+ years. Furthermore, many traits are highly influenced by the particular growing environment or correlated with conditions such as seed tuber size.

Many of the key traits are qualitative, controlled by a few genes, or are quantitative in nature but highly influenced by major genes. Yet, because potato is an autotetraploid, five distinct genotypes may be observed at any given locus (AAAA, AAAa, AAaa, Aaaa, aaaa) making progeny testing difficult.

Ideal Markers

There are various types of molecular markers, with Single Nucleotide Polymorphisms (SNPs) being the most widely used type today.

Ideally, the markers are:

- Co-dominant, revealing the detail of allelic composition
- “Perfect” (i.e. located within the gene itself, to avoid the chance of recombination between the marker and the tagged gene)
- Consistently reproducible (i.e. the assay results are consistent for each genotype, even across different laboratories)
- High-throughput (HTP) (i.e. amenable to automation for fast delivery of results and widespread application)
- Cost effective.

In potato, many single-gene traits have been mapped in addition to major genes contributing to the expression of quantitative traits.

Potential Advantages to Exploit with MAS

MAS may offer some significant advantages over phenotypic selection.

For example:

1. Easier screening (e.g. resistance screening for potato cyst nematode)
2. Screening at an earlier plant development stage
3. Greater accuracy by eliminating classification errors due to environmental effects
4. Selection can be applied outside of the market region and growing season (e.g. in an off-season nursery or greenhouse)
5. HTP operation facilitates larger population sizes and higher selection intensity
6. Fewer years of testing
7. May be less expensive than phenotypic evaluation

Potential cost savings are just as important in optimizing the product pipeline as time savings, and depending on the situation at hand, may be more important.

Slater et al. (2013) compared the cost of applying MAS to screen for disease resistance in the second field generation (G_2) against the cost of phenotypic evaluation of G_2 selections. The MAS option was 37.3% of the cost of the phenotypic selection option. This strategy also provides the option to select for other markers associated with single-gene traits or major genes among G_2 clones, with multiplexing marker assays as means to further the cost savings. In addition to cost savings, the MAS option resulted in a one-year time savings in cultivar development.

See Slater et al. 2013 for more detail.

Extracting the Most Value From Deployment of MAS

Often, the advantages provided through technology can be amplified further by tailoring the product pipeline for optimal fit or combining technologies.

Slater et al. (2014) exploited the advantages of MAS further in potato by combining MAS with selection based on estimated breeding value (EBV). By evaluating G_2 clones for EBV, those individuals with high scores that may not possess the disease resistance sought through MAS are not discarded but recycled to increase the frequency of favorable alleles for other target traits. Those individuals selected by MAS and EBV are advanced to regional trials, reducing testing by two years. This enhances the time savings achieved overall to three years.

Although Slater et al. (2014) utilized pedigree information to compute EBVs, this could also be done through the use of markers using an approach like genomic selection.

See Slater et al. 2014 for more detail.

Boosting Breeding Efficiency Through Doubled Haploidy

With many crops, once a cross has been made to develop a breeding population, a major time lag occurs between F_2 and the development of a fully inbred line. This time lag can be mitigated through early testing, yet we have seen that phenotypic measurement is most accurate and precise when non-segregating materials are used.

Doubled haploidy is a technology that can be used to by-pass the process of inbreeding, providing a quick route to homozygosity with a high degree of fidelity. With doubled haploid (DH) production systems available for more than 250 plant species, this technology manipulates the double fertilization process which gives rise to seed with a 2N embryo and a 3N endosperm (Fig. 6).

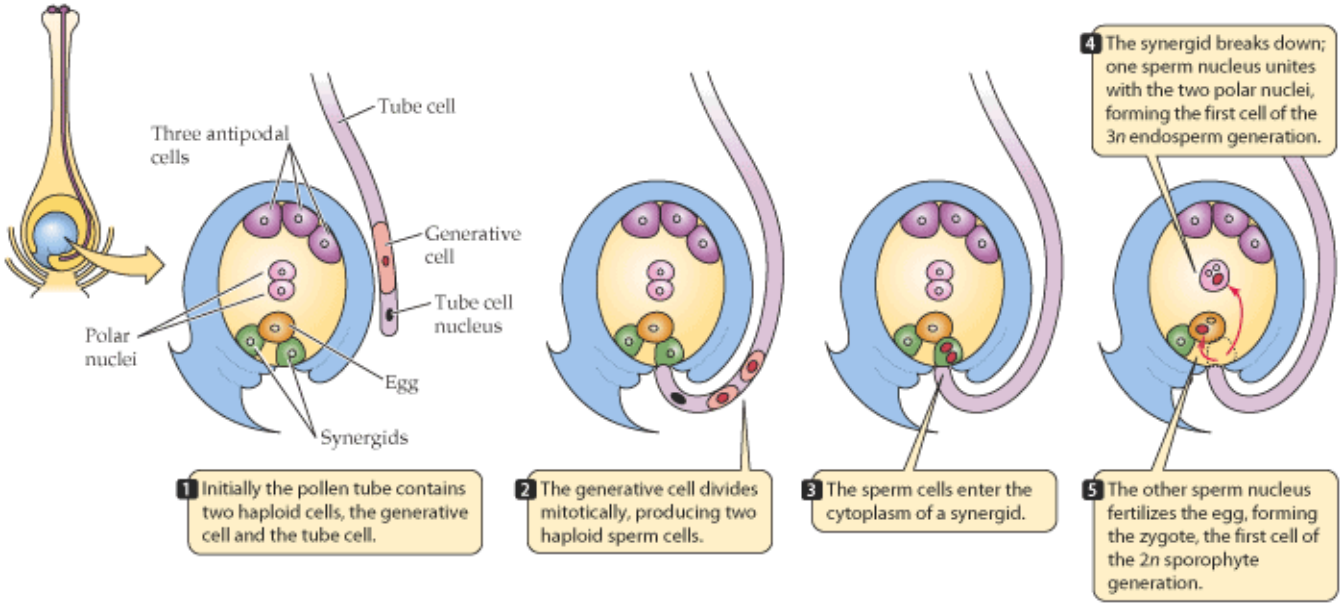


Fig 6. The process of double fertilization. Licensed CC BY 4.0

In maize, the use of an *in vivo* maternal haploid system via **gynogenesis** has become common. First reported by Stadler and Randolph in 1929, the system involves the use of an "inducer line" and typically also features a color marker system to facilitate the identification and confirmation of haploid individuals.

In vivo Maternal DH Production Process

The basic steps involved in *in vivo* maternal DH production in maize are (Fig. 7):

1. Pollinate **donor** plants (F1 or F2 progeny from a breeding cross) with inducer line

2. Identify and recover haploid seed via color marker phenotype
3. Germinate haploid seed and apply chromosome doubling agent
4. Grow DHs to maturity & self-pollinate

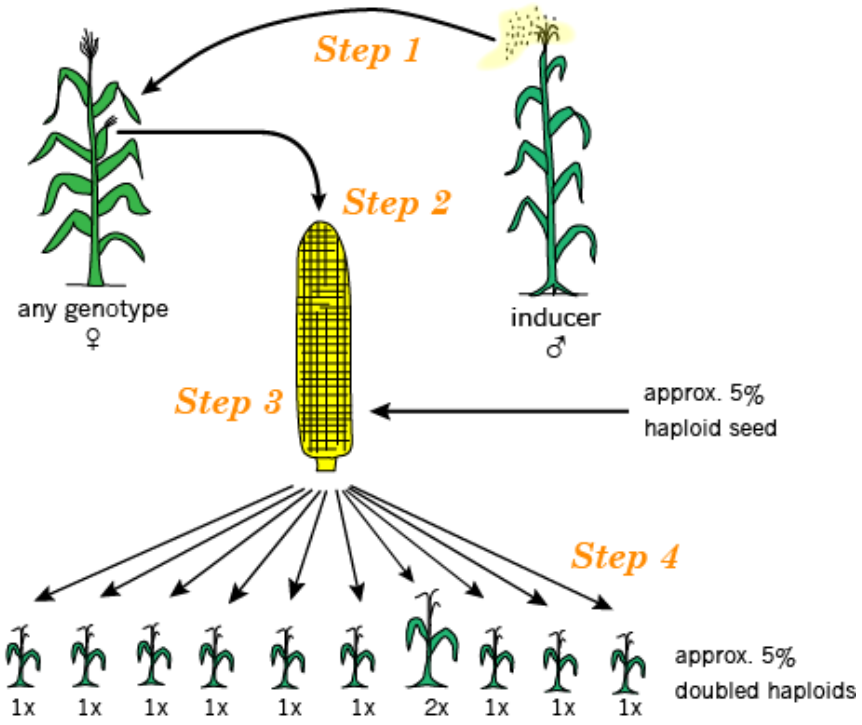


Fig. 7 Steps in *in vivo* maternal doubled haploidy.

Haploid Selection

Much of the seed produced by pollinating the donor plants with the inducer line will be F_1s, which are useless. A small proportion will be haploids which contain only genetic material from the donor line (the haploid induction rate is determined by the inducer line). And a fraction will be lethal mutants. There also may be some seed resulting from self-pollination or outcrossing since pollinations are made manually.

The color marker system makes it possible to sort the various types of seed. Since the color indicates genetic contribution from the donor, the haploid seed will show color only in the endosperm, not the scutellum which harbors the embryo (Fig. 8).

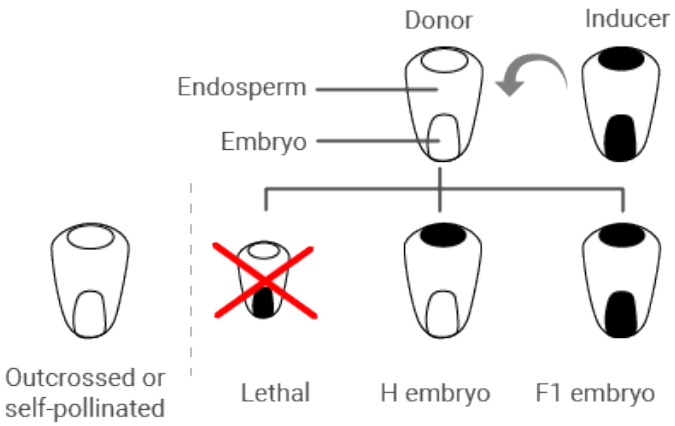


Fig. 8 Possible outcomes of use of a haploid inducer line in maize based on color marker system.

In real life, expression of the color markings is influenced by the genetic background (Fig. 9).

Fig. 9 From left to right: haploid, diploid, and outcross kernels are shown. Photo courtesy of Eunsoo Choe, University of Illinois.

Chromosome Doubling Treatment

Once haploid seeds are recovered, these are germinated and then exposed to a chromosome-doubling agent to duplicate the chromosome content.

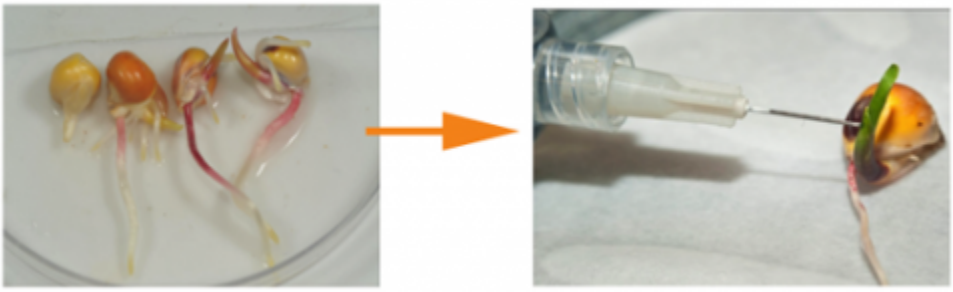


Fig. 10 Injecting germinated kernels with colchicine. Photos courtesy of Eunsoo Choe, University of Illinois.

Recovering Doubled Haploid Lines

Those haploid plants that successfully doubled chromosome content are self-pollinated and grown to physiological maturity. The harvested selfed seed represents the new doubled haploid line. All seed on the ear is genetically identical and completely homozygous.

For more details on the *in vivo* maternal doubled haploid process in maize, see the YouTube clip from CIMMYT: "Doubled Haploids: A simple method to improve the efficiency of maize breeding."

Extracting the Maximal Value from Deployment of Doubled Haploidy

Doubled haploidy offers the benefit of reduced cycle time. Plus, testing homozygous lines results in more precise estimates of performance than segregating lines, and thus affords more accuracy in selection decisions. Used to estimate QTL effects, DH lines again win over segregating lines.

Some organizations offer DH services so this step in the breeding process may be outsourced if desired. Nonetheless, there is an investment in tapping into this resource which can improve efficiency in cultivar development.

The big question is how to extract maximal value from the deployment of this technology. We will focus on three ways:

- The type of donor population
- Pre-selection of lines for haploid induction, sometimes referred to as "F_2 enrichment"
- The structure of performance testing given the vast number of DHs that can be produced from a single breeding cross.

Best Population Type to Serve as Donor

Some options for the type of donor population to use to initiate doubled haploidy include F_1, F_2, and BC_1. These population types differ regarding the number of meiosis, that is, the number of times cell division has taken place to produce reproductive cells. With meiosis comes recombination, and after all, it is new combinations of favorable alleles that we seek in developing improved cultivars!

Genetic theory tells us that the number of crossover events is less among DH lines compared to RILs. The evidence supporting this includes:

- Smith et al. (2008) found a mean of 15 recombinations per RIL versus 10 recombinations per DH genome developed from F_1 donors
- Smith et al. (2008) found that the percentage of lines with ≥4 intact parental chromosomes (10 chromosomes in maize) was 37% among DHs and 13% among RILs.

DH lines developed from F_2 donors allow for one more meiosis to occur. Work by Bernardo (2009) showed that DH F_2 lines were no more than 3% lower in selection response than RILs and were up to 6% higher

in selection response than DH F_1 lines. Likewise, BC_1 donors would offer the same advantage as F_2 donors (compared to F_1).

Although one more generation is required to produce the F_2 or BC_1 donors (compared to the F_1 donors), it does not necessarily mean another year is required in development time. Off-season nurseries or greenhouses could be utilized to advance some/all generations and facilitate more generations per year. So this is a good trade-off!

Pre-Selection of Lines for Haploid Induction

Using F_2 (or BC_1) lines provides an opportunity for some selection to be performed before pollination with the inducer line. The rationale for such selection is to focus DH resources on individuals with greater genetic promise. Better donors should translate to greater genetic gain toward the product target.

Selection among F_2 families based on single-plant performance could focus on high-heritability traits or marker genotype. Examples include:

- Plant morphology
- Disease resistance
- Favorable marker genotype for linked trait
- Against negative marker genotype for linked trait
- Increased or decreased frequency of recombinants in which repulsion linkages have been broken or favorable complexes of gene combinations have been preserved (Smith et al., 2008).

In a more elaborate scheme, Bernardo et al. (2010) proposed three cycles of genomic selection in a recurrent selection scheme as a means to enrich the frequency of favorable alleles for yield and other quantitative traits of interest.

The advantage of pre-selection of lines is to focus breeding and testing resources on lines with the greatest genetic potential, which should in turn result in the greatest genetic gain and higher R.

Performance Testing and Selection of DH Lines

With doubled haploidy, it is easy to produce a large number of progeny per breeding cross. *And* this can be accomplished in an accelerated timeframe (compared to inbreeding).

Through doubled haploidy, the breeder's problem of having to wait until RILs are developed is solved. *Now* the issue is how to test the vast number of quickly-available inbred lines! An example of how to address this issue is briefly described below.

Practical Examples

Assuming similar budgets Melchinger et al (2005), compared the rate of genetic gain (G) achieved with a **four-location** preliminary yield test with a population of **250 segregating testcross progeny** in maize with two-stage testing involving **739 DH lines** at one location, followed by testing of **29 selections** (from Year 1) at **nine locations** in a second season. The two-stage testing resulted in a nearly 20% increase in G with the same resource investment.

Although the number of yield plots is the same in both cases, two-stage testing accommodates more lines and focuses a higher degree of scrutiny on the top performers. Thus, two-stage preliminary testing may be more effective in maximizing R when doubled haploid technology is utilized to develop progeny.

Summary

To optimize the product pipeline and deliver improved cultivars to farmers most effectively and efficiently, consider each aspect of process design in breeding and testing as it relates to the factors influencing response to selection R and rate of genetic gain (G).

Technologies can be useful in boosting efficiency. The merit of applying any new technology can be judged in terms of its ability to increase genetic gain and the cost/benefit ratio. Once a decision is made to incorporate a new technology into the product pipeline, its value can be maximized by fine-tuning the pipeline process to take advantage of opportunities the technology affords to increase R and reduce development time.

Finally, integration of all the components of the process is critical to a robust product pipeline that consistently produces improved cultivars. A chain is only as strong as its weakest link!

References

Allen, F.L., R.E. Comstock, and D.C. Rasmusson. 1978. Optimal environments for yield testing. Crop Sci. 18: 747-751.

Bernardo, R. 2009. Should maize doubled haploids be induced among F_1 or F_2 plants? Theor. Appl. Genet. 119:255-262.

Bernardo, R. 2010. Breeding for Quantitative Traits in Plants. 2nd edition. Stemma Press, Woodbury, MN.

Bernardo, R., R.E. Lorenzana, and P.J. Mayor. 2010. Exploiting both doubled haploids and cheap and abundant molecular markers in corn breeding. pp 40-53. In Proc. 46th Annual Corn Breeders' School, 2010.

Falconer, D.S. 1989. Introduction to Quantitative Genetics, 3rd ed. John Wiley & Sons, New York, NY.

Fehr, W.R. 1991. Principles of Cultivar Improvement v1: Theory and Technique. Macmillan Publishing Company.

Gauch, H.G. 1992. Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs. Elsevier, Amsterdam, The Netherlands.

Geiger, H.H. 2009. Doubled haploids. pp641-657. In J.L. Bennetzen and S. Hake (eds), Handbook of Maize: Genetics and Genomics. Springer, New York, NY. https://doi.org/10.1007/978-0-387-77863-1_32

Melchinger, A.E., C. F. Longin, H.F. Utz, J.C. Reif. 2005. Hybrid maize breeding with doubled haploid lines: quantitative genetic and selection theory for optimum allocation of resources. Pp 8-21. In Proc. 41th Annual Corn Breeders' School, March 2005.

Mumm, R.H. 2013. A look at seed product development with genetically modified crops: Examples from maize. J. Agricultural and Food Chemistry 61(35): 8254-8259. DOI 10.1021/jf400685y

Randolph, L.F. 1932. Some effects of high temperature on polyploidy and other variations in maize. Genetics 18: 222-229.

Slater, A.T., N.O.I. Cogan, J.W. Forster. 2013. Cost analysis of the application of marker-assisted selection in potato breeding. Molecular Breeding 32: 299-310. DOI 10.1007/s11032-013-9871-7

Slater, A.T., N.O.I. Cogan, B.J. Hayes, L. Schultz, M.F.B. Dale, G.J. Bryan, J.W Forster. 2014. Improving breeding efficiency in potato using molecular and quantitative genetics. Theor. Appl. Genet. 127: 2279-2292. DOI 10.1007/s00122-014-2386-8

Smith, J.S.C., T. Hussain, E.S, Jones, G. Graham, D. Podlich, S. Wall, M. Williams. 2008. Use of doubled haploids in maize breeding: Implications for intellectual property protection and genetic diversity in hybrid crops. Molecular Breeding 22: 51. https://doi.org/10.1007/s11032-007-9155-1

Yan, W., J.B. Holland. 2010. A heritability-adjusted GGE biplot for test environment evaluation. Euphytica 171: 355-369. DOI 10.1007/s10681-009-0030-5

How to cite this chapter: Mumm, R.H. (2023). Optimization of Product Pipeline. In W. P. Suza, & K. R. Lamkey (Eds.), *Cultivar Development*. Iowa State University Digital Press.

Chapter 7: Launching Improved Cultivars

Rita H. Mumm

Preparing For New Product Release

Once a superior line meeting the product target is identified, and it is determined that the potential new cultivar can be reliably reproduced, steps toward product release are set in motion. The preparations for launching a new cultivar are largely a function of the Supply Chain, which is the final stage of the process of cultivar development (Fig. 1).

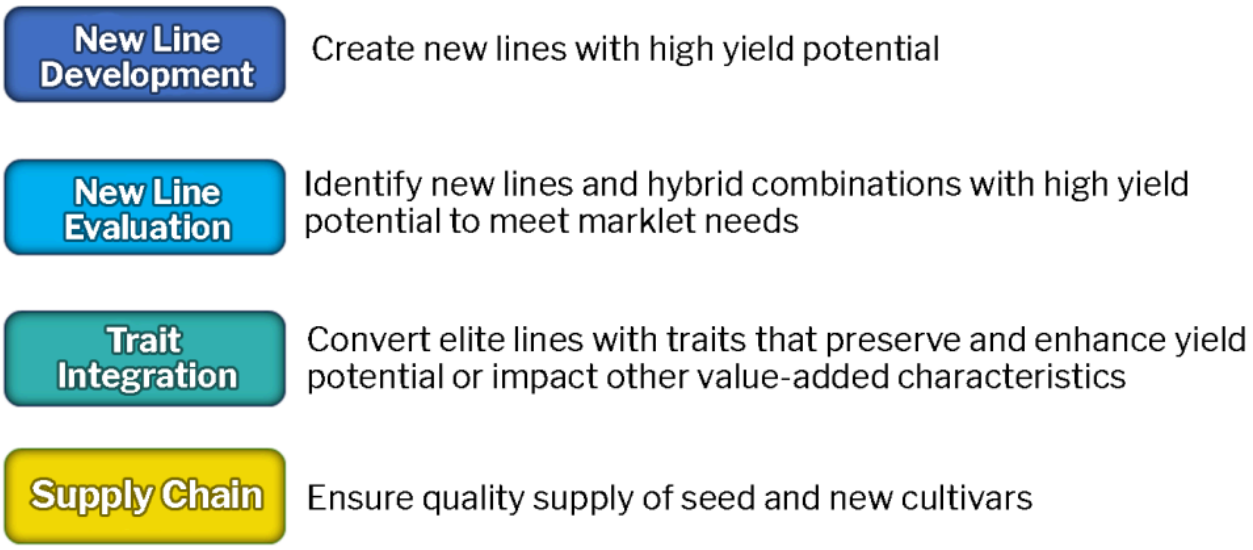


Fig. 1 The four core functions in the product pipeline.

Variety Registration and Varietal Performance Testing

The new cultivar must represent a new genetic entity (not previously commercialized); it must be distinguishable from other varieties. In many countries, **variety registration testing**, often referred to as **DUS**, is conducted to establish the identity of the new cultivar based on four aspects:

- *Distinctiveness* from other existing varieties.
- *Uniformity* such that individual plants reflect consistency in quality and in the level of variation within the population.
- *Stability* of distinctive characteristics manifests through generations of propagation.
- *Varietal Identity* in terms of defining morphological characteristics.

In DUS testing, the performance of the potential new cultivar is compared to a wide range of existing varieties

to validate its unique benefits and attributes. **A detailed description of DUS testing guidelines and procedures is available from the** International Union for the Protection of New Varieties of Plants (UPOV).

In addition, **Varietal Performance Testing**, often referred to as **VCU**, is conducted to demonstrate the merit of the new cultivar to stakeholders in the value chain, i.e., producers, end-users, and consumers. Typically, the test results must prove the new cultivar is superior to existing varieties in one or more aspects, e.g., yield, disease resistance, and nutritional value. VCU is typically conducted through on-farm trials for multiple years.

The new cultivars that meet requirements for registration and performance through DUS and VCU testing are officially approved for commercialization and are listed in a variety register. Note that requirements vary from country to country. DUS and VCU requirements need to be accommodated in the design of the testing regime for cultivar development.

Priorities in Scale-up to Commercial Quantities

To launch the new cultivar, that is, release it for commercial distribution, it is critical to maintain genetic integrity, and purity, and to keep the propagation materials free from pests and diseases as the scale-up to commercial quantities is performed.

Critical issues differ depending on whether the new cultivar is seed-propagated or vegetatively propagated.

Clean Stock of Clonally-propagated Cultivars

To generate a “clean” stock of a clonally-propagated cultivar, a heat treatment may be applied ahead of isolating meristematic tissue for culturing to kill off any pathogens. Clean plant stock source materials can be stored long-term through cryopreservation (Fig. 2).

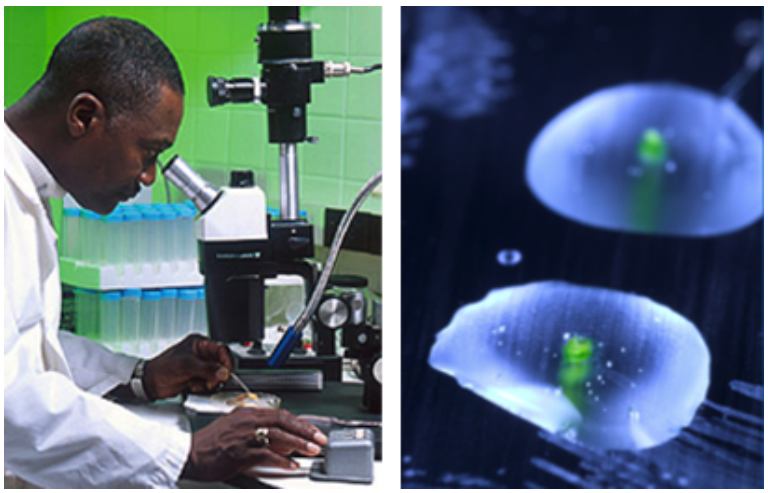


Fig. 2 Tissue isolation for culturing (left) and cryopreserved shoot tips (right). Photos courtesy of USDA ARS.

To maintain and increase clean stock, best practices include:

- Physical isolation, i.e., planting distance from potential sources of infection.
- Control of insect vectors of disease through the use of screen houses or cages, chemical controls, or distance/time isolation.
- Inspection and testing, e.g., visual observation, use of indicator plants/grafting, and/or diagnostics such as ELISA assays or molecular marker analysis to monitor for pathogen presence.

Producing Commercial Volumes of Seed-propagated Cultivars

For new cultivars propagated by seed, seed quality is a key priority at every stage in seed production and scale-up (Fig. 3).

Seed quality reflects a number of parameters to ensure genetic integrity and purity, seed vigor, the absence of pathogens, and cleanliness. Because producing commercial volumes of the new cultivar is a multi-step process, seed quality must be safeguarded through each step to achieve the desired result for distribution to farmers.

Fig. 3 Small quantity (left) and large volume (right) of maize seed.

Seed Classes From Breeder's Seed to Commercialization

It all starts with **breeder's seed**, which is produced in the initiating breeding program through controlled hand pollination. Breeder's seed represents the standard of genetic identity to which all other increases of the line will be compared. Breeder's seed is provided to Supply Chain through a hand-off from the plant breeder who developed the new cultivar. The development and maintenance of breeder seed is an important obligation of the plant breeder.

The breeder's seed is used by the Foundation group to increase the variety or parent line, producing pre-basic and foundation seed. In the case of a hybrid cultivar, Hybrid Seed Production follows to create the hybrid seed that will be sold to farmers/growers and, in the case of 3-way or double-cross hybrid cultivars, any intermediary hybrid materials. See Chapter 3 for more on the types of hybrids and requirements for utilizing hybridization in crop improvement (e.g., male sterility systems).

Certified seed production is governed by the national seed regulations of the country in which the seed is to be produced, with the aim to produce seed for commercial distribution that meets standards for genetic purity, seed germination, seed moisture, and the presence of premium value-added traits. Internationally recognized, the Organization for Economic Co-operation and Development (OECD) provides guidelines for seed certification (Table 1). (Note that certified seed is not required for commercial distribution by every country for every crop.)

Table 1 Seed classes in commercial seed production, particularly in countries following OECD Certification Standards.

Seed Class	Produced from	Comprised of
Pre-basic seed	Breeder seed	Cultivar *per se* -or- parent line of hybrid cultivar
Foundation seed (or Basic seed)	Pre-basic or breeder's seed	Cultivar *per se* -or- parent line of hybrid cultivar
Registered seed	Foundation seed	Progeny of foundation seed or registered seed used to produce certified seed
Hybrid seed	Crosses between parent lines increased by Foundation seed group	Hybrid cultivar -or- intermediary hybrid in the case of 3- or 4-way hybrid cultivars
Certified seed	Foundation or higher seed classes	Cultivar *per se* including hybrid cultivar

See "Molecular Plant Breeding, Chapter 12. Genomic Tools for Variety Registration and Protection", for more detail regarding rules for certified seed production.

Best Practices in Seed Production

With the goal to produce quality seed for cultivar release, steps to safeguard genetic purity are implemented at each stage of seed production. Logistics are arranged to ensure the designated pollen and seed parents are involved and to maximize seed returns.

Safeguards include:

- Proper isolation of production fields to minimize the chance of stray pollen. The physical isolation distance per crop depends on pollen mobility. For maize, an isolation distance of at least 300 meters is typically employed. Fields can also be isolated in time; other fields in close proximity are arranged so that pollen shed is not occurring at the time when production fields are receptive to pollen. Furthermore, other fields in proximity to the seed production field are monitored throughout the season to check for risks of pollen migration.
- Field selection to facilitate control of soil-borne disease (no previous history) and absence of volunteer plants (of the same species).
- Ideal agronomic management (e.g., fertility, planting, disease control, soil moisture, etc.) to facilitate top yields.
- Careful cleaning of planting and harvest equipment (e.g., planters, combines, threshers, trucks, bins, etc.) to ensure no seed carryover and prevent inadvertent contamination through seed mixing.
- Judicious placement of insect pollinators (as needed depending on the crop) to facilitate pollination.

- Field inspection throughout the growing season.
- Careful record keeping to maintain chain of custody documentation from seed source through each seed increase to final production.
- Accurate labeling of harvested seed to safeguard genetic authenticity.

Additional Best Practices For Production of Hybrid Seed

Fig. 4 Hybrid production of maize. Photo courtesy of Iowa State University.

In hybrid production, care is taken in the placement of the female parent plants to the male parent plants to ensure ample movement of pollen from male rows to female rows. In addition, the ratio of male rows to female rows is designed to facilitate sufficient pollination. Corn hybrid production typically features a 2:1 ratio as seen in the photo (Fig. 4); four rows of female inbred flanked by two rows of male inbred.

With some crops such as rice, movement of the relatively heavy pollen grain is aided by mechanical means. For example, a helicopter flown overhead at peak pollen shed will move pollen by wind to female plants.

It is critical that male and female parent plants are flowering simultaneously. This is referred to as "**nicking**." Sometimes to achieve the nick, one parent is planted later than the other in keeping with known maturity of the lines. With some crops, the earlier female or male parent may be "trimmed" to trigger a second flowering.

Generally, male rows in a hybrid production field are harvested or removed from the field ahead of harvesting of the female rows to prevent contamination of the hybrid seed (with selfed seed of the male parent).

Foundation Research

Once the hand-off of breeder seed to supply chain occurs, it is the responsibility of the Foundation seed group to maintain the genetic integrity of the variety or parent seed. With seed increases, measures are taken to prevent genetic drift, contamination through pollen migration, and unintended seed mixing, and to ensure accurate seed labeling (as described previously).

In addition, the Foundation seed group often includes a research arm to optimize seed multiplication and hybrid production. Each line it maintains is characterized in depth and, for hybrid cultivars, parent lines are assessed to facilitate best practices to achieve optimal seed quality and output in hybrid production in a cost-effective manner.

Seed Quality Tests

Once harvested, each **lot** of seed of the new cultivar undergoes screening and testing to determine seed quality and purity.

Evaluations may include:

- The ratio of seed to other materials such as weed seed and inert matter.
- Genetic authenticity to ensure proper labeling.
- Seed purity and percentage of off-types.
- Rate of germination as a measurement of seed vigor.
- Seed moisture content to indicate whether further drying of the seed is required to prolong shelf life.
- Freedom from adventitious presence of GM events.
- Freedom from disease and pests.

Genetic Authenticity and Seed Purity

Genetic authenticity (and seed purity) may be verified through **DNA fingerprinting** to ensure the genetic makeup of the seed matches that of the intended cultivar. This DNA fingerprint also represents a genetic profile to support intellectual property protection of the new cultivar. See "Molecular Plant Breeding, Chapter 12. Genomic Tools for Variety Registration and Protection" for details on establishing the DNA profile for varietal registration as well as protection of new cultivars.

To monitor seed purity, a "**grow out**" of each seed lot is typically performed to check for characteristic morphological properties, absence of off-types, and percentage of seed exhibiting any premium value-added traits (e.g., herbicide tolerance, oil profile). In this way, the genetic purity of the given seed lot can be determined as well as the extent to which the seed lot sample conforms to the prescribed standard.

For the grow-out, the size of the sample representing a given seed lot depends on the predetermined threshold for off-types. For example, if no more than 1 percent off-types are allowed, at least 400 seeds (i.e., plants) must be tested. For stricter standards, larger sample sizes are required (Table 2).

Table 2 Sample size required for various threshold percentages for off-types and minimum levels of genetic purity.

Threshold for off-types (%)	Minimum level of genetic purity	Sample size per seed lot
0.01	99.9	4000
0.20	99.8	2000
0.30	99.7	1350
0.50	99.5	800
1.00	99.0	400

The grow-out should be conducted in the target market region in a field free from potential volunteer plants. Furthermore, the grow-out must be managed in keeping with agronomic practices recommended to farmers for production of the new cultivar including planting date, plant spacing, fertilization regime, and other cultural practices.

The percentage of off-types is estimated for each seed lot from the results of the grow out and the number of seeds (i.e., plants) in the sample which do not conform to the prescribed standard.

Seed Germination Testing

To determine the germination rate of a seed lot, a sample of seed from each seed lot is collected, cleaned, and germinated on a moist substrate under laboratory conditions, with replication. The mean percentage of seeds that germinate to produce normal seedlings can then be estimated for that seed lot (Fig. 5).

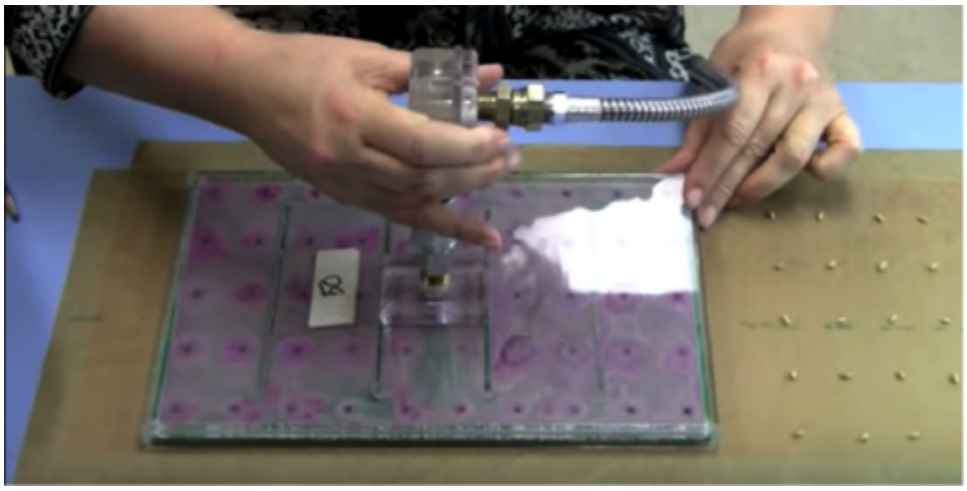


Fig. 5 Testing of seed germination. Photo courtesy of University of California at Davis.

Phytosanitary Testing

Seed lots must be checked for the presence of seed-borne disease and insect pests and disease-carrying organisms (Fig. 6). Such tests serve as a backup to field scouting during seed increase and hybrid production.

Fig. 6 Corn seed tested for *Diplodia* (left) and *Setaria* seed soak test for nematodes (right). Photos courtesy of C.C. Block and J.W. Van Roekel, USDA ARS.

Testing for Adventitious Presence of GM Events

Just as seed lots are evaluated for the percentage of plants expressing premium value-added traits, seed may also evaluated for the absence of other traits, particularly those resulting from genetically modified (GM) events, that are not intended as part of the genetic package of the cultivar.

Because traits created through genetic modification (e.g., GM events) are subject to regulation on a country-by-country basis, seed must be tested before product release to ensure that only desired GM events, consistent with seed labeling, are present. This type of testing, which evaluates for **adventitious presence**, is conducted through DNA analysis, using markers for the DNA sequence characteristic of known events.

See "Molecular Plant Breeding, Chapter 12. Genomic Tools for Variety Registration and Protection" for more detail on monitoring for absence or presence of transgenes.

Overview of Production of a Given Seed Lot

In summary, an effective seed production process utilizes best practices and features points at which monitoring is conducted to verify seed integrity, purity, and quality. The graphic below (Fig. 7) depicts examples of test points that might be implemented in a seed increase or hybrid seed production.

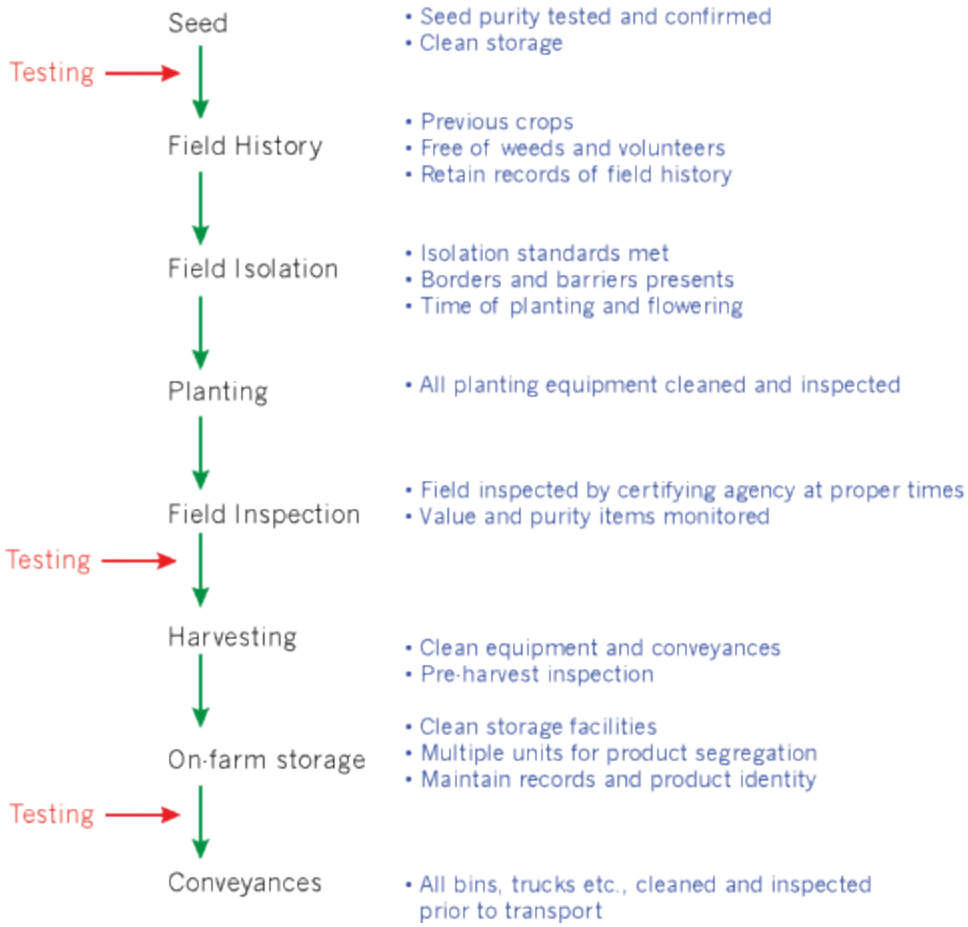


Fig. 7 Test point examples. Adapted from Sundstrom et al., 2002.

Packaging Seed for Distribution

Seed lots that pass the strict standards for various tests are stored under conditions conducive to long seed life, controlling moisture/temperature levels.

To prepare seed for distribution of the new cultivar:

- seed is cleaned to remove any weed seed or inert matter.
- seed treatments may be applied as desired (e.g., seed-applied fungicides or microbials to stimulate seedling growth).
- seed is packaged to prevent absorption of water from the atmosphere and to keep seed contained and inaccessible to insects and disease.
- seed bags/containers are labeled.

The seed container must be accurately labeled (or “tagged”) (Fig. 8), listing information related to the specific variety, the seed lot and origin, seed testing results of the seed lot, seed purity, etc. Varieties featuring value-

added traits typically list the thresholds of seed purity for these traits. With certified seed, tags reveal these important details and more.

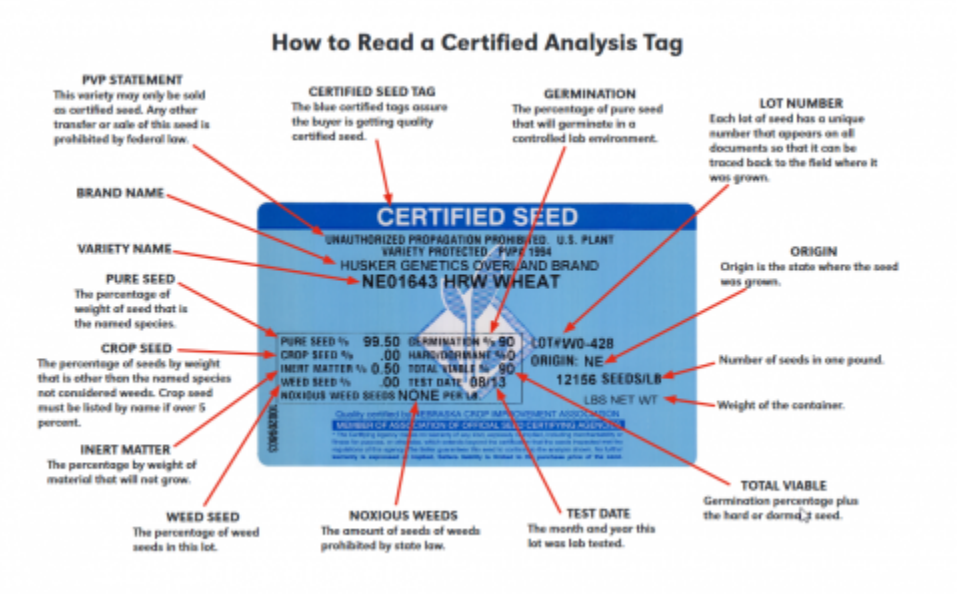


Fig. 8 Certified Analysis Tags contain details on each seed lot. From CROPWATCH. Courtesy of the University of Nebraska-Lincoln.

Promoting the New Cultivar

The breeder has a role in creating awareness and demand for the new cultivar. He/she may set up demonstration plots to highlight soon-to-be-released varieties or new value-added traits to farmers (Fig. 9). Field days at the research station represent a good opportunity for breeders to exhibit characteristics of new varieties or traits to farmers and growers.

Furthermore, the breeder can make recommendations as to appropriate agronomic practices to maximize and protect yields, which are critical to farmers' success, especially with new cultivars.

The goal is to promote rapid adoption of new, improved cultivars.

Fig. 9 Field days. Photos courtesy of CIMMYT (left) and Oregon State University (right).

Protecting Intellectual Property

The new, improved cultivar represents an important investment of skills, time, energy, and resources. To recoup the investment and to provide the incentive for further development of improved varieties that meet stakeholder needs, the intellectual property represented by the new cultivar must be protected.

A number of mechanisms are available to protect the interests of plant breeders and foster the development of a robust national seed sector, including:

- Plant Variety Protection (PVP) for seed and tubers.
- Plant patents for asexually propagated plants, except tubers.
- Utility patents for any type of plant showing a particular utility or purpose.
- Contracts.
- Trade secrets.

Plant Variety Protection (PVP)

A key mechanism recognized worldwide is Plant Variety Protection (PVP), which provides intellectual property protection for unique varieties of a sexually-reproduced plant or tuber-propagated plant for 20 years (25 years for trees and vines).

PVP is governed by the International Union for the Protection of New Varieties of Plants (UPOV), which was established in 1961 as an outcome of the International Convention for the Protection of New Varieties of Plants. The mission of UPOV is to provide and promote an effective system of plant variety protection, which encourages the development of new, improved varieties of plants for the benefit of society. PVP provides the basis for plant breeders' rights to intellectual property pertaining to their innovation.

To apply for PVP protection, the breeder of the new cultivar (or his/her employer) must file an application with the authorities of UPOV. The African Intellectual Property Organization operates a plant breeders' rights system which covers the territory of 17 member states. Applications can be filed with authorities entrusted with the responsibility of granting breeders' rights.

PVP Provisions

PVP is not intended to restrict farmers from saving seed for their own purposes, nor does it restrict private use of the variety for non-commercial or experimental purposes. PVP does allow the use of the protected variety in breeding to develop other improved cultivars, provided the new cultivar is not **essentially derived** from the protected cultivar. Thus, the protected variety can still contribute to the advancement of improved varieties during the 20 years of coverage; PVP does not slow innovation.

PVP does guard against marketing and selling of the protected variety, or a variety essentially derived from the protected variety, by other commercial entities. PVP also forbids multiplication of the protected variety or import/export without the express authorization of the breeder.

However, enforcement of the PVP falls to the breeder and his/her organization. As a basis for such enforcement, a DNA fingerprint of the protected variety produced in accord with established standards of protection in the crop species can be used to expose the genetic identity of suspected offenders.

Once the term of PVP protection has expired, the variety is considered "public domain." Thus, the net result is to encourage more innovation, which ultimately promotes food security and benefits the consumer.

Sharing Germplasm

Plant breeders are advised to keep good records of germplasm in inventory as well as germplasm shared with others. Such records will be important in protecting intellectual property embodied by new cultivars.

A **Standard Material Transfer Agreement** (**SMTA**) is a contract between the provider and the recipient of the germplasm. Basically, the provider agrees to share the germplasm and other available non-confidential descriptive information about the plant materials. The recipient agrees to use the materials for research, breeding, or agricultural training, without claim to intellectual property. Should the recipient commercialize a product that incorporates or is developed from the material, the SMTA outlines expectations for compensation, if any, to the provider.

See Crop Improvement: Genetic Variation and Germplasm Usage for more detail on SMTAs.

Starting the Cycle Anew

This course in Cultivar Development has focused on the cycle of cultivar improvement and design of the process pipeline to effectively facilitate the efficient development of improved crop cultivars.

The launch of a new, improved cultivar that meets the needs of stakeholders embodied in the product target represents the close of one cycle and triggers the start of another. A plant breeder's job never ends!

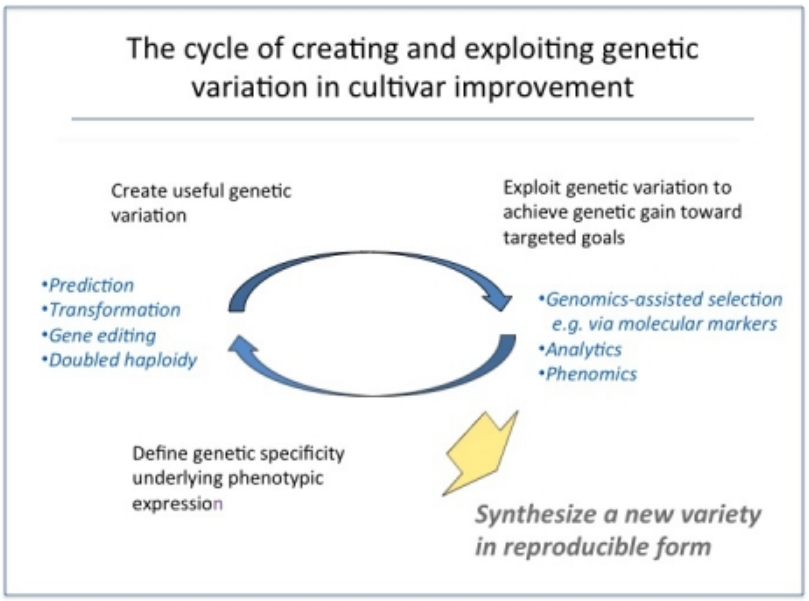


Fig. 10 Inclusion of modern tools in the cycle of creating and exploiting genetic variation in cultivar improvement. Adapted from Mumm, 2013.

Now, with the start of a new cycle, product targets are redefined or adjusted in keeping with reassessment of stakeholder needs and desires. Armed with the information gleaned from past cycles, the breeder is better informed with respect to genetic architecture of traits of interest and how such traits interact with the environment, and each other, and is better able to choose parents to create new breeding populations, identify or create effective testing environments, and employ technologies (Fig. 10). Thus, knowledge gleaned can lead to better ways to create useful genetic variation and more effective ways to exploit this variation to achieve stated product targets.

What types of information acquired during the development of an improved cultivar might serve to better guide choice of parents and to enhance evaluation/selection of progeny in future cycles?

References

Baenziger, P.S., R.H. Mumm, R. Bernardo, E.C. Brummer, P. Langridge, P. Simon, and S. Smith. 2017. Plant Breeding and Genetics—A paper in the series on "The Need for Agricultural Innovation to Sustainably Feed the World by 2050." Council for Agricultural Science and Technology (CAST), Issue Paper 57, Ames, Iowa.

Bishaw, Z., and A.J.G. van Gastel. 2009. Variety release and policy options. In S. Ceccarelli, E.P. Guimaraes, E. Weltizien (eds), Plant Breeding and Farmer Participation. Food and Agriculture Organization of the United Nations (FAO), Rome. FAO

MacRobert, J.F., P.S. Setimela, J. Gethi, and M.W. Regasa. 2014. Maize Hybrid Seed Production Manual. Mexico, D.F.: CIMMYT.

Mumm, R.H. 2013. A look at seed product development with genetically modified crops: Examples from maize. J. Agricultural and Food Chemistry 61(35): 8254-8259. DOI 10.1021/jf400685y

Mumm, R.H., and D.S. Walters. 2001. Quality control in the development of transgenic crop seed products. Crop Sci. 41:1381-1389.

Sundstrom, F.J., J. Williams, A. Van Deynze, and K.J. Bradford. 2002. Identity preservation of agricultural commodities. University of California, Agricultural Biotechnology in California Series, ANR Publication No. 8077. ANR Catalog, UC-Davis

USDA ARS National Laboratory for Genetic Resources Preservation. Clonal Brochure No. 11182016. USDA-ARS.

Xu, Y. 2010. Intellectual property rights and plant variety protection. pp 512-549. In Molecular Plant Breeding. CAB International, Cambridge, MA.

How to cite this chapter: Mumm, R.H. (2023). Launching Improved Cultivars. In W. P. Suza, & K. R. Lamkey (Eds.), *Cultivar Development*. Iowa State University Digital Press.

Contributors

Editors

Walter Suza

Suza is an Adjunct Associate Professor at Iowa State University. He teaches courses on Genetics and Crop Physiology in the Department of Agronomy. In addition to co-developing courses for the ISU Distance MS in Plant Breeding Program, Suza also served as the director of Plant Breeding e-Learning in Africa Program (PBEA) for 8 years. With PBEA, Suza helped provide access to open educational resources on topics related to the genetic improvement of crops. His research is on the metabolism and physiology of plant sterols. Suza holds a Ph.D. in the plant sciences area (with emphasis in molecular physiology) from the University of Nebraska-Lincoln.

Kendall Lamkey

Lamkey is the Associate Dean for Facilities and Operations for the College of Agriculture and Life Sciences at Iowa State University. He works in collaboration with the dean, associate deans, department chairs, college-level centers, and other unit leaders to ensure that operations directly advance the mission of the college and that resources are deployed wisely and efficiently. Previously, he served as the chair for the Department of Agronomy at Iowa State University, where, in addition to advocating for research and the PBEA program, he oversaw the Agronomy Department's educational direction, its faculty, and Agronomy Extension and Outreach. Dr. Lamkey is a corn breeder and quantitative geneticist and conducts research on the quantitative genetics of selection response, inbreeding depression, and heterosis. He holds a Ph.D. in plant breeding from Iowa State University and a master's in plant breeding from the University of Illinois. Lamkey is a fellow of the American Society of Agronomy and the Crop Science Society of America and has served as an associate editor, technical editor, and editor for *Crop Science*.

Chapter Authors

Rita H. Mumm

Contributors

Anthony A. Mahama, Gretchen Anderson, Todd Hartnell, Andy Rohrback, Tyler Price, Glenn Wiedenhoeft, and Abbey K. Elder

Applied Learning Activities

The following downloadable Applied Learning Activities (ALAs) and recommended readings are associated with the chapters linked below:

Chapter 1

Learning Activities

- FAO Database Exercise [PDF]
- Changes in organizational infrastructure [PDF]
- Example Commercial Soybean Improvement Program [PDF]

Recommended Readings

- Eathington SR, Crosbie TM, Edwards MD, Reiter RS, Bull JK (2007) Molecular markers in a commercial breeding program. Crop Science 47(S3) S154–S163.
- Moose SP, Mumm RH (2008) Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiology 147: 969–977.
- Pannar 2017 Product Catalogue [PDF]

Chapter 2

Learning Activities

- Example Soybean Product Target [PDF]
- Graphic of population types [PDF]
- BMS Crop Ontology Curation Tool [PDF]
- Breeding Plan Template [DOC]

Recommended Readings

- Knapp SJ (1998) Marker–assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Science 38: 1164–1174.
- Mikel MA, Diers BW, Nelson RL, Smith HH (2010) Genetic diversity and agronomic improvement of North American soybean germplasm. Crop Science 50: 1219–1229.

- Mumm RH, Dudley JW (1994) A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. Crop Science 34:842-851.

Chapter 3

Learning Activities

- Exercise in analysis of GCA and SCA [PDF]
- Example Hybrid Maize Product Target [PDF]
- Example Commercial Hybrid Maize Improvement Program [PDF]
- Breeding Plan Template_Blank [DOC]

Recommended Readings

- Eberhardt SA. 1971. Regional maize diallels with US and semi–exotic varieties. Crop Science 11: 911–914.
- Dudley JW. 1987. Modification of methods for identifying inbred lines useful for improving parents of elite single crosses. Crop Science 27: 944–947.
- Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Australian J. Biol. Sci. 9: 463–493
- Pannar 2017 Product Catalogue [PDF]

Chapter 4

Learning Activities

- Product Target Template_Blank [DOC]
- Example Commercial Potato Cultivar Improvement Program [PDF]

Recommended Readings

- Gruneberg W. Mwanga R, Andrade M, and J Espinoza. 2009. Selection Methods: Breeding clonally propagated crops. In Plant Breeding and Farmer Participation, S Ceccarelli, EP Guimaraes, E Weltzien (eds), Food and Agriculture Organization of the United Nations, Rome, Italy.
- Bradshaw, J.E., and M. Bonierbale. 2010. Potatoes. In Handbook of Plant Breeding: Root and Tuber Crops, J.E. Bradshaw (ed). Springer Science+Business Media, LLC, New York, NY, USA

Chapter 5

Learning Activities

- Breeding Plan Template_Blank [PDF]
- Table from Sedcole 1977 [DOC]

Recommended Readings

- Peng, T., X. Sun, and R.H. Mumm. 2014a. Optimized breeding strategies for multiple trait integration: I. Minimizing linkage drag in single event introgression. Molecular Breeding 33:89–104. DOI 10.1007/s11032–013–9936–7.
- Peng, T., X. Sun, and R.H. Mumm. 2014b. Optimized breeding strategies for multiple trait integration: II. Process efficiency in event pyramiding and trait fixation. Molecular Breeding 33:105–115. DOI 10.1007/s11032–013–
9937–6.

Chapter 6

Learning Activities

- P i table from Falconer 1989 [PDF]
- Example Commercial Potato Cultivar Improvement Program [PDF]

Recommended Readings

- J. Sánchez–Martín, D. Rubiales, F. Flores, A.A. Emeran, M.J.Y. Shtaya, J.C. Sillero, M.B. Allagui, E. Prats (2014). Adaptation of oat (Avena sativa) cultivars to autumn sowings in Mediterranean environments. Field Crops Research. 156: 111–122.
- Slater, A.T., Cogan, N.O.I. & Forster, J.W. Cost analysis of the application of marker–assisted selection in potato breeding. Mol Breeding 32, 299–310 (2013). https://doi.org/10.1007/s11032–013–9871–7
- Slater, A.T., Cogan, N.O.I., Hayes, B.J. et al. Improving breeding efficiency in potato using molecular and quantitative genetics. Theor Appl Genet 127, 2279–2292 (2014). https://doi.org/10.1007/s00122–014–2386–8
- Watson A, Ghosh S, Williams MJ, Cuddy WS, Simmonds J, Rey M–D, Asyraf Md Hatta M, Hinchliffe A, Steed A, Reynolds D, Adamski NM, Breakspear A, Korolev A, Rayner T, Dixon LE, Riaz A, Martin W, Ryan M, Edwards D, Batley J, Raman H, Carter J, Rogers C, Domoney C, Moore G, Harwood W, Nicholson P, Dieters MJ, DeLacy IH, Zhou J, Uauy C, Boden SA, Park RF, Wulff BBH, Hickey LT. 2018. Speed breeding is a powerful tool to accelerate crop research and breeding. Nature Plants doi:10.1038/s41477–017–0083–8.

- Weyhrich, R.A., Lamkey, K.R. and Hallauer, A.R. (1998), Responses to Seven Methods of Recurrent Selection in the BS11 Maize Population. Crop Science, 38: 308–321. https://doi.org/10.2135/cropsci1998.0011183X003800020005x

Chapter 7

Learning Activities

- CIMMYT Standard Material Transfer Agreement [PDF]