Published on *Plant Breeding E-Learning in Africa* (https://pbea.agron.iastate.edu)

Home > Course Materials > Molecular Plant Breeding > Genome Construction

Genome Construction

By Thomas Lübberstedt, Walter Suza, William Beavis (ISU)

Introduction

Introduction

One of the main challenges in plant breeding is development of best marker assisted breeding method for complex traits. At the present, marker-based approaches are limited in their ability to detect and quantify marker-trait relationships, in particular for traits that are under the influence of gene x gene and gene x environment interactions. Also, as you have learned in previous lessons of this course, QTL estimates are biased by population size and limited set of environments, making QTL estimates less suitable for crop improvement. For this reason, simulation modeling is an emerging important tool to choose among proposed breeding methods because experimental evaluation of breeding methods is time and resource limited. Another challenge is management of multiple breeding objectives for several complex traits, making it more likely that an operations research approach called multi-objective optimization will gain favor in crop breeding. Thus, this lesson will introduce operations research as a tool to address multiple crop breeding objectives.

Fig. 1 Operations research is a tool to address crop breeding objectives. Photo by Iowa State University.

Objectives

- Recapitulate the concept of genetic gain
- Introduce the concept of multi-objective optimization
- Introduce the concept of operations research in plant breeding

Fig. 2 Evaluating materials requires expert training, patience, equipment and time. Photo by Iowa State University.

Recapitulation of the Concept of Genetic Gain

Definition

Genetic gain (∆G) is defined as predicted change in the mean value of a trait within a population as a result of selection. The ∆G equation (Fig. 3) allows comparison of predicted effectiveness of particular breeding methods and helps breeders decide how resources should be allocated for achieving various breeding objectives.

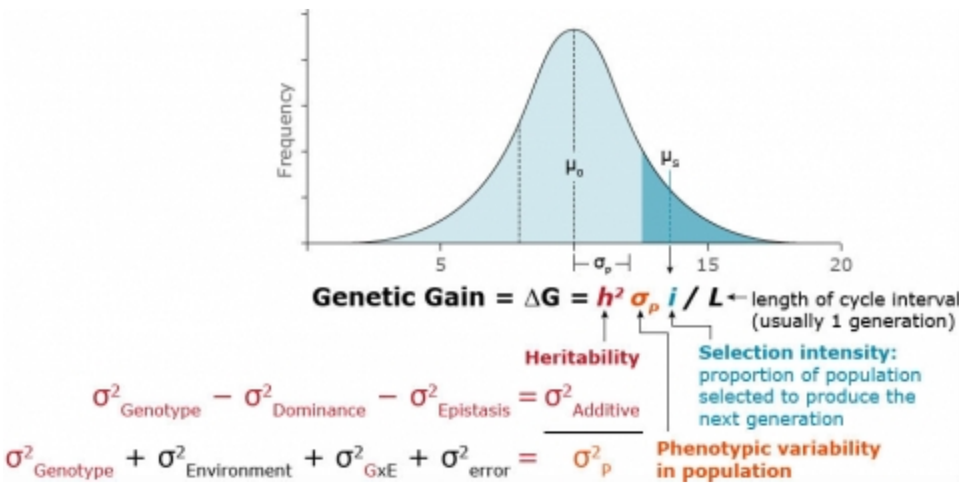


Fig. 3 The genetic gain equation and its components. The curve illustrates distribution showing frequency of individuals in breeding population (y axis) that display various phenotypic values (x axis) of individuals in a breeding population. For the above curve, the mean phenotypic value of the original population is denoted μ_0, and the mean phenotypic value for the selected individuals is denoted μ_S. Genetic components (σ^2) and phenotypic distribution (σ_P) are indicated. Adapted from Moose and Mumm, 2008.

Commercialization Challenges

Figure 4 illustrates a generic plant breeding program involving mating, evaluation, selection, and testing of breeding materials resulting in commercialization of a cultivar. Such a program faces the challenges of time to commercialization of a cultivar, and resources allocated to obtain such cultivar from thousands of individuals.

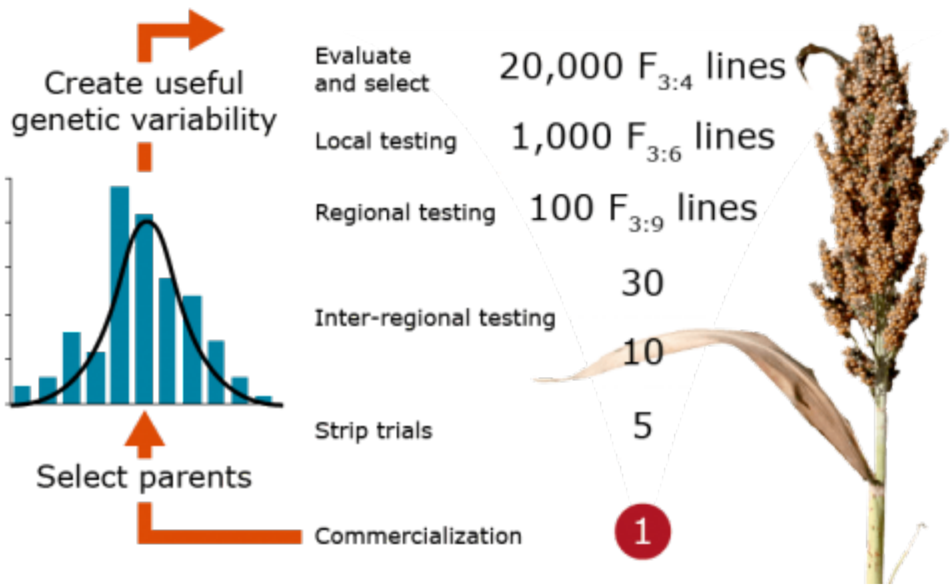


Fig. 4 An example of a breeding program

Crop Yield Progress

Despite such challenges, from the 1940s, the yields of corn and soybean in the United States have continued to rise (Fig. 5) mainly due to improvement in crop genetics and agronomic practices.

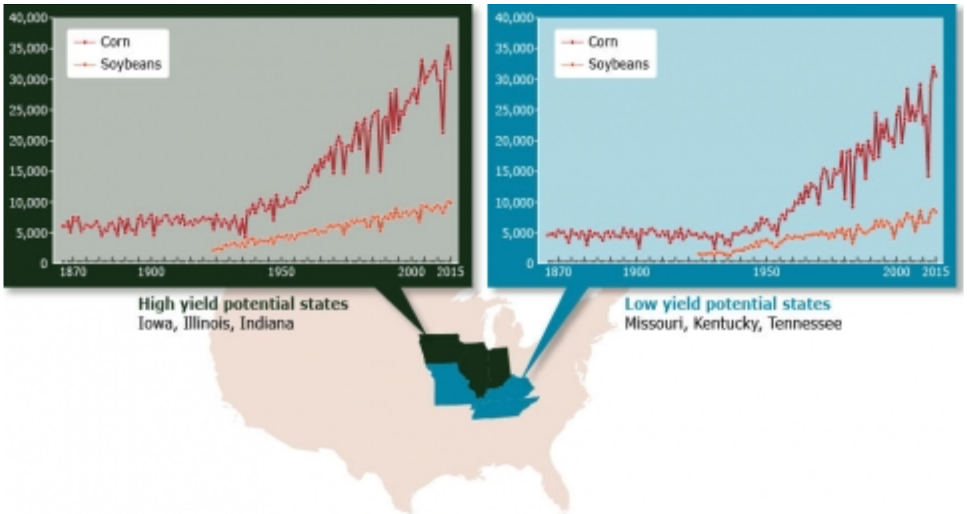


Fig. 5 Average corn and soybean yields (in kg/ha) for the U.S. from 1866 to 2015 in high-yield potential states (Iowa, Illinois, and Indiana) and low-yield potential states (Kentucky, Missouri, and Tennessee). Data from National Agricultural Statistics Service.

Global Food Demand Trends

Despite the upward trend in crop yields in the US and other parts of the world, rising human and animal populations will pose a greater demand for more to be produced per unit of land. The growing global demand for food (Fig. 6) raises the question of whether it is possible to double the current level of production in the next 20 years (Fig. 7). Undoubtedly, to reach 300 bushels/acre of corn by 2030 will require cutting-edge approaches in genomics and breeding. But the problem will be the cost of reaching such a high level of yield with limited time and resources. Thus, integration of new approaches, for example, Genomic Selection, transgenics, and operations research, may be necessary. The next lesson sections entail application of operations research tools in plant breeding as a novel approach to increase ΔG.

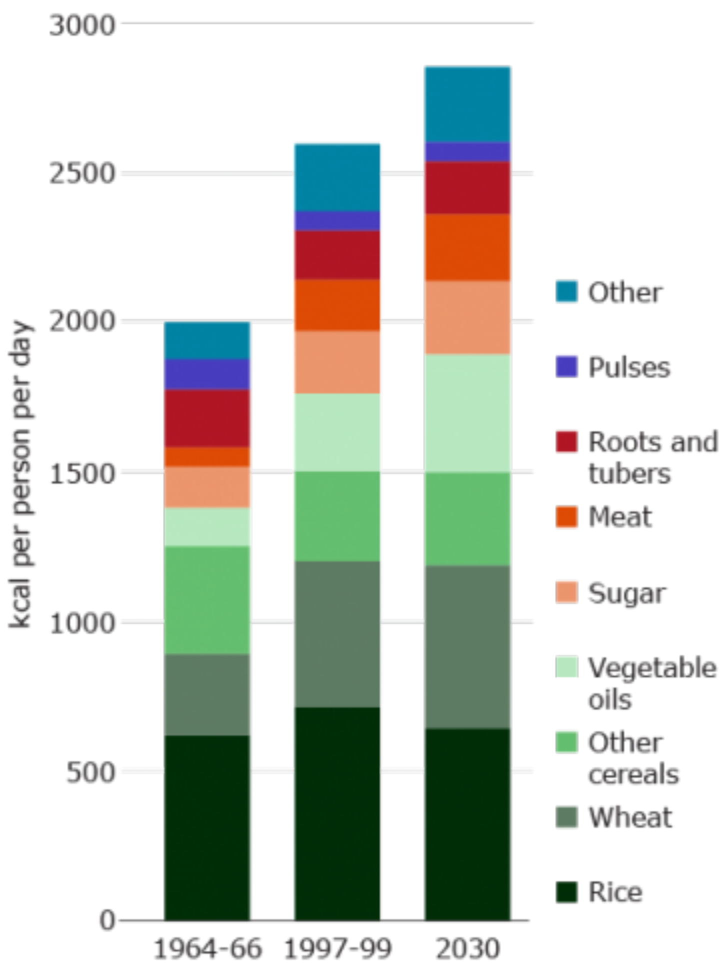


Fig. 6 Global progress in food consumption. Adapted from FAO, 2002.

Need for Advancement

Historically plant breeding has been a form of art: to create new varieties. Thus, ∆G has depended on management of resources, to produce new varieties; while optimization has been ignored. Nonetheless, plant breeding has the potential to become an engineering discipline, relying on operations research, which will be necessary for average yields to double by 2030 (Fig. 7).

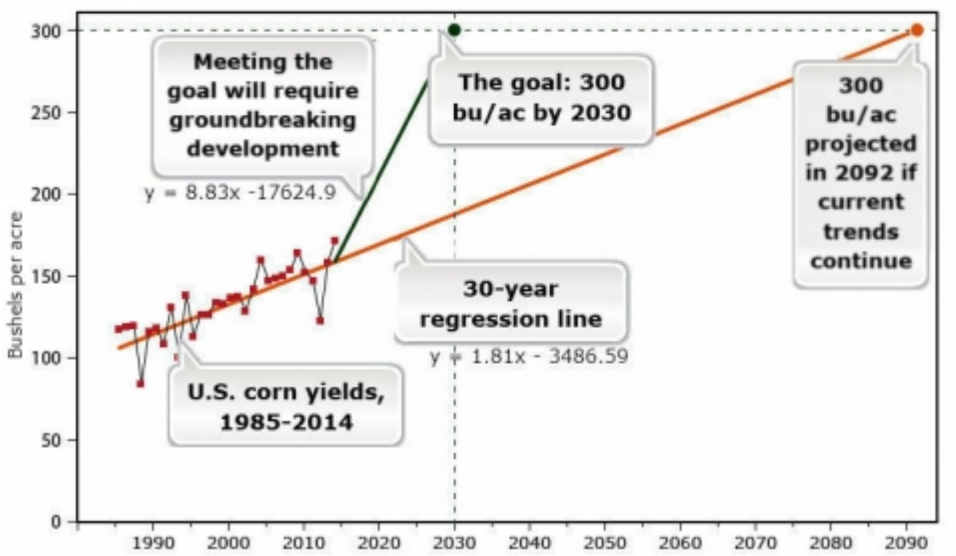


Fig. 7 New plant breeding tools will be needed to produce 300 bushels (Bu) per acre by 2030. Data from National Agricultural Statistics Service.

Multi-Objective Optimization

Introduction

Multi-objective optimization (MO) is an operations research approach used in various fields, including engineering, finance, biomedicine and management. Optimization involves application of more than one objective processes for evaluation that can take into account multiple criteria that need to be considered for making a decision. Therefore, as information on plant genomes continue to emerge, it is now possible to apply the MO approach for large scale plant breeding (Xu et al., 2011). For example, a plant breeding goal may have two objectives, 1) selection and fixation of desirable genes at a set of loci controlling a trait of interest, and 2) keeping genetic variability at the remaining loci to retain adaptability. The challenge of applying MO to solve these competing objectives is identification of optimal solutions to the problems (Chinchuluun and Pardalos, 2007). Such solutions are called Pareto optimal solutions, and they are a measure of MO optimization efficiency.

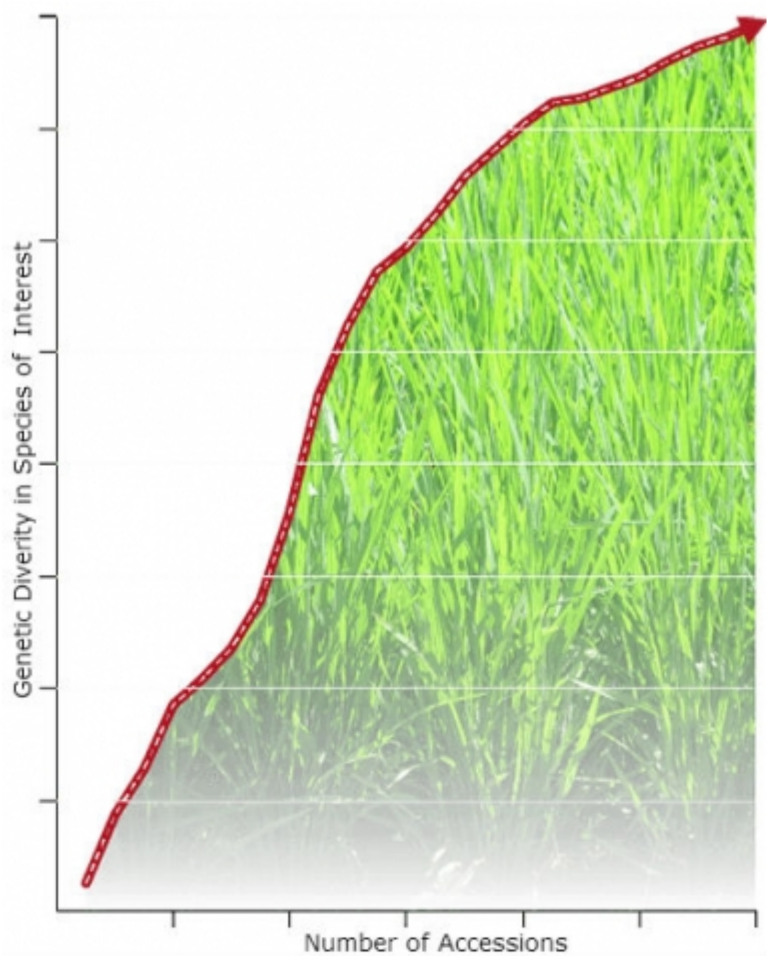


Fig. 8 Some multi-objective optimization problems involve maximizing one variable while minimizing another.

Pareto optimal solutions

We will not dwell on the mathematics used to derive Pareto optimal solutions in this lesson. But it is important to know that there usually exist multiple Pareto optimal solutions for MO problems, and searching for all Pareto optimal solutions can be expensive and time consuming (Chinchuluun and Pardalos, 2007). Nonetheless, recent advances in computational research suggest that it is possible to obtain Pareto optimal solutions for plant breeding problems within reasonable computation time (Xu et al., 2011). Such solutions will be useful tools to help plant breeders make informed decisions in the world of large amounts of genomics data for multiple breeding objectives for complex traits. 2011). Such solutions will be useful tools to help plant breeders make informed decisions in the world of large amounts of genomics data for multiple breeding objectives for complex traits.

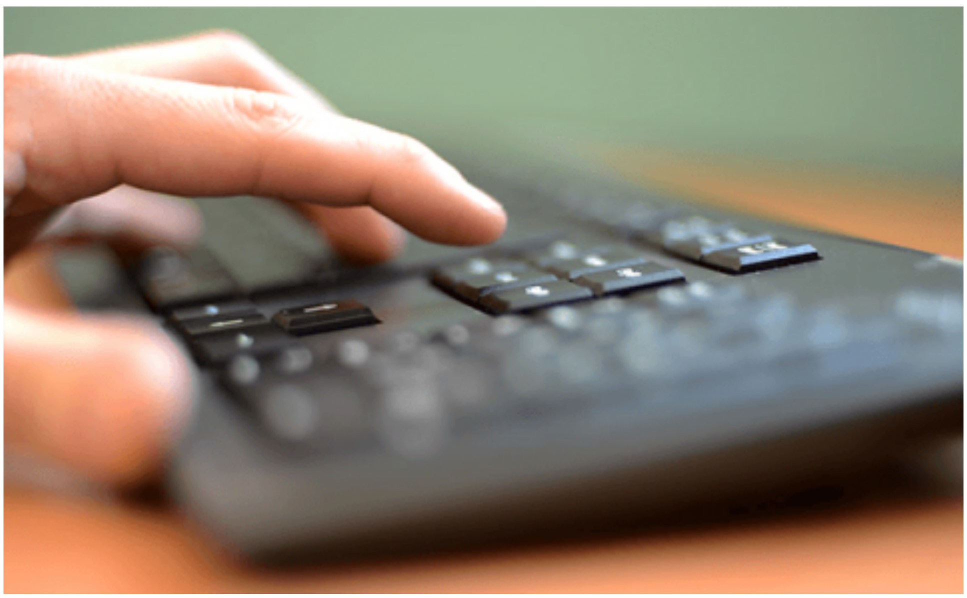


Fig. 9 Computer systems can now yield Pareto optimal solutions in acceptable time frames.

Operations Research in Plant Breeding

Introduction

Operations Research involves the application of mathematical models to provide optimal solutions to a problem. An OR approach (Fig. 10) consists four components, 1) Problem, 2), Model, 3) Algorithm, and 4) Solver.

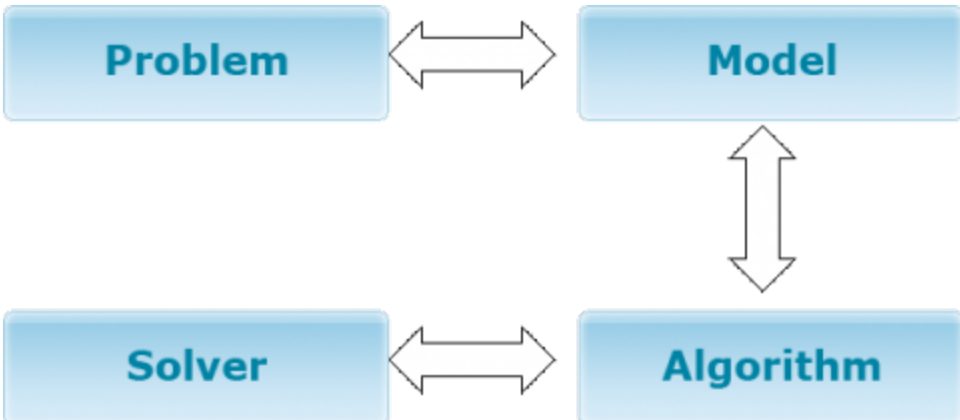


Fig. 10 A multi-objective optimization plant breeding problem requires the use of optimization models, algorithms, and computer technologies. Complexity of the problem, robustness of model, algorithm used, and computer solvers influence the cost of solving the problem.

Step 1: Defining the Problem

There is an original population of individuals (Fig. 11). Each individual has a pair of chromosomes, and each chromosome has a number of genes. Some genes are undesirable, while the desirable ones have different variants. The desirable genes will be assigned a value of 1 and undesirable 0. What is the best way to assemble all variants of desirable genes into a target population?

Original population

1	2	3	4	5	6	7	8	9	10
0	0	0	0	1	0	0	0	0	1
C	A	A	C	B	C	C	A	A	A
1	0	1	0	0	0	0	1	0	1
B	C	B	A	C	B	A	A	B	B
1	1	0	1	0	1	0	1	0	0
C	A	B	A	C	B	A	A	C	B
0	0	1	0	1	1	0	0	0	0
B	C	C	C	A	B	C	B	C	C
0	0	0	0	1	0	0	0	1	1
C	A	A	B	A	B	C	A	B	B

Integer Programming →

Min (G) = 2

Target population

T1	T2	T3	T4
1	1	1	1
A	A	B	C
1	1	1	1
B	C	A	B
1	1	1	1
B	A	C	C
1	1	1	1
C	B	A	A
1	1	1	1
B	A	A	C

Fig. 11 Operations research can help assess the possibility of stacking genes into multiple backgrounds.

Step 2: Developing a Model

A model has four key elements - data, decisions, objective, and constraints.

A. Data

B. Decisions

A decision would have to be made about number of data and recombination points, and number of chromosomes in the target population.

C. Objective

The objective is to maximize probability of getting the target population.

D. Constraints

Constraints can be, for example, the number of chromosomes in target population without undesirable alleles, but such that all desirable variants are retained. Also, the maximum number of recombination events could be another constraint.

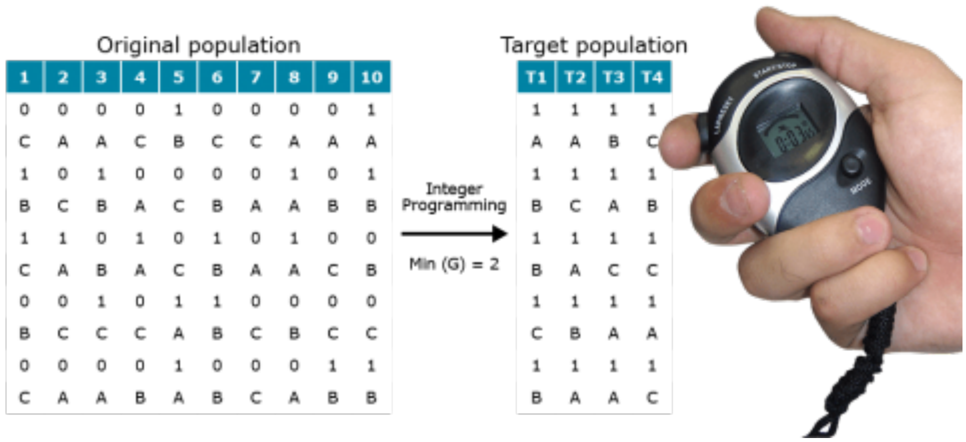
Step 3: Designing a Suitable Algorithm

The problem in this example belongs to a class of so-called non-deterministic polynomial-time hard (NP-hard) problems (Xu et al. 2011). Importantly, if an algorithm solves one NP-hard problem, it can be used to solve all other NP-hard problems.

Fig. 12 Often a computer program will be used as an algorithm for solving NP-hard problems, such this one for the Traveling Salesman Problem.

Step 4: Solving the Problem

Computation time spent to solve the problem in Figure 11 was 0.03 seconds (W.D. Beavis, personal communication).



Genome Construction vs. Genomic Selection

The hypothesis is that genome construction is better than genomic selection (Fig. 13). The hypothesis is developed from the premise that a target genotype can be defined. However, if the target genotype has to be determined using experimental methods, then GS will be more effective because experimental methods are underpowered and biased.

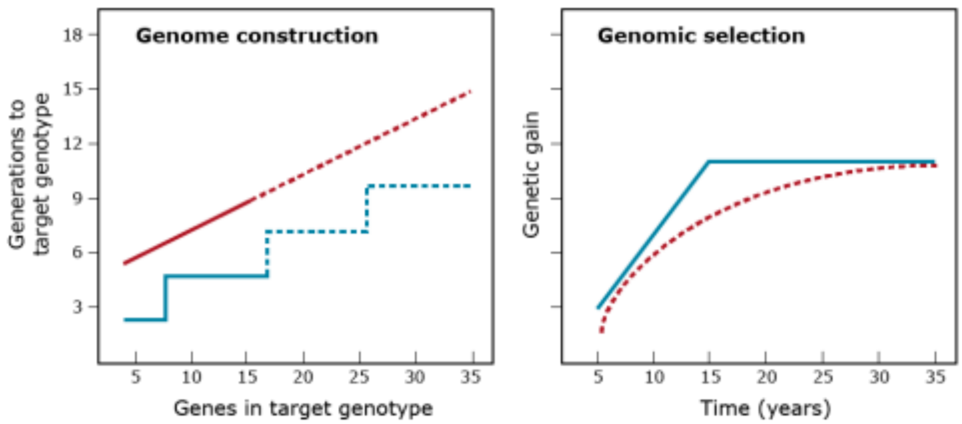


Fig. 13 Comparison of genome construction and genomic selection approaches

Reflection

The Module Reflection appears as the last "task" in each module. The purpose of the Reflection is to enhance your learning and information retention. The questions are designed to help you reflect on the module and obtain instructor feedback on your learning. Submit your answers to the following questions to your instructor.

1. In your own words, write a short summary (< 150 words) for this module.
2. What is the most valuable concept that you learned from the module? Why is this concept valuable to you?
3. What concepts in the module are still unclear/the least clear to you?

References

Chinchuluun, A. P. M. Pardalos. 2007. A survey of recent developments in multiobjective optimization. Ann. Oper. Res. 154: 29-50. http://link.springer.com/content/pdf/10.1007%2Fs10479-007-0186-0

Egli, D. B. 2008. Comparison of corn and soybean yields in the United States: Historical trends and future prospects. Agron J. 100: S-79-S-88. https://www.agronomy.org/publications/aj/abstracts/100/Supplement_3/S-79

Expert Meeting on "How to Feed the World in 2050," FAO, Rome, 24-26 June 2009. http://www.fao.org/docrep/012/ak542e/ak542e00.htm

FAO, 2002. World agriculture: towards 2015/2030. United Nations, 2002.

Moose, S. P., and R. H. Mumm. 2008. Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiol. 147: 969-977. http://www.plantphysiol.org/content/147/3/969.full.pdf+html

Xu, P., L. Wang, and W. D. Beavis. 2011. An optimization approach to gene stacking. Europ. J. Oper. Res. 214: 168-178. http://ac.els-cdn.com/S0377221711003559/1-s2.0-S0377221711003559-main.pdf?_tid=d56559dc-8cdc-11e2-9ffb-00000aacb35d&acdnat=1363289267_62b034810a8c0fc45567932c53ce45a3

Acknowledgements

This module was developed as part of the Bill & Melinda Gates Foundation Contract No. 24576 for Plant Breeding E-Learning in Africa.

Molecular Plant Breeding Genome Construction Author: Thomas Lübberstedt, Walter Suza, and William Beavis (ISU)

Multimedia Developers: Gretchen Anderson, Todd Hartnell, and Andy Rohrback (ISU)

How to cite this module: Lübberstedt, T., W. Suza, and W. Beavis. 2016. Genome Construction. *In* Molecular Plant Breeding, interactive e-learning courseware. Plant Breeding E-Learning in Africa. Retrieved from **https://pbea.agron.iastate.edu**.