

How do “conventional” and “advanced” breeding methods compare?

Lecture 7

of

Genetic improvement of crops

CJ Yang

Nov 1st, 2022

Breeding methods

Conventional breeding

- Pedigree
- Single seed descent (SSD)
- Doubled haploid (DH)
- Bulk
- Backcross
- Population/evolution
- Hybrid
- Clonal
- Mutation
- Reciprocal selection

Covered in lecture 5

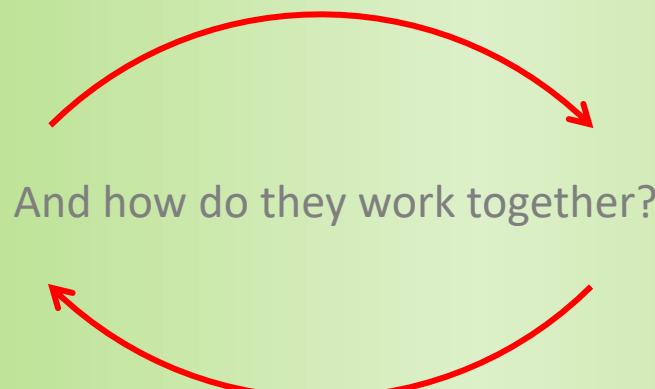
How do they compare?

Advanced breeding

- Marker assisted selection (MAS)
- Genomic selection (GS)
- Genome wide association (GWAS)
- Linkage mapping
- Genetic modification (GM)
- Gene editing (GE)
- Omics: phenomics, enviromics, shovelingomics, ...
- Speed breeding (= SSD?)

Covered in lecture 6

And how do they work together?



Metrics to consider in our comparison

1. Time and cost
2. Available infrastructure/technology
3. Current knowledge
4. Genetic diversity
5. Bandwagon and hype



Conventional is better



Advanced is better



Time and cost

Varies according to crop, trait and size of breeding program.

Field/greenhouse trial

- Plot size, maintenance, phenotype, harvest, ...
- E.g. strawberry vs wheat



<https://www.countrysideonline.co.uk/hobbies-and-leisure/food-and-cooking/extending-the-strawberry-season/>



<https://www.bbc.co.uk/news/uk-scotland-49594680>

Genotyping

- KASP ~ \$0.40/sample/marker
- Array ~ £40/sample/40k markers (Barley, Darrier et al 2019)
- GBS ~ £60/sample/40k markers (Barley, Darrier et al 2019)
- MAS = KASP, GS/GWAS/Linkage mapping = Array/GBS.

GM/GE

- Expensive, often focused on few genes.
- Trials for GM can be complicated and costlier.

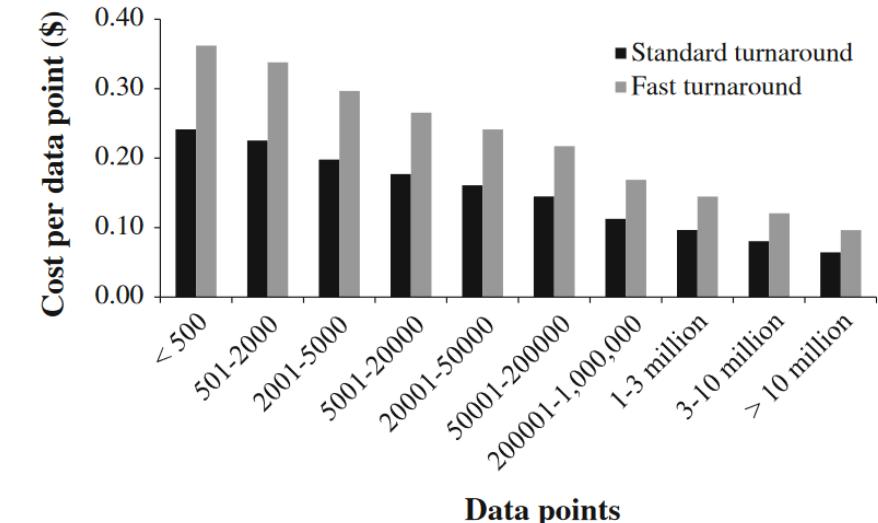


Fig. 4 KASP genotyping cost per data points for standard (5–8-week) and fast (2–3-week) data turnaround. This genotyping cost is specific to the Generation Challenge Program molecular breeding platform users, with CIMMYT being the major user of the service, and may not be the same for others

Semagn et al (2013)

Time and cost (an example in potato breeding)

Cost analysis of the application of marker-assisted selection in potato breeding

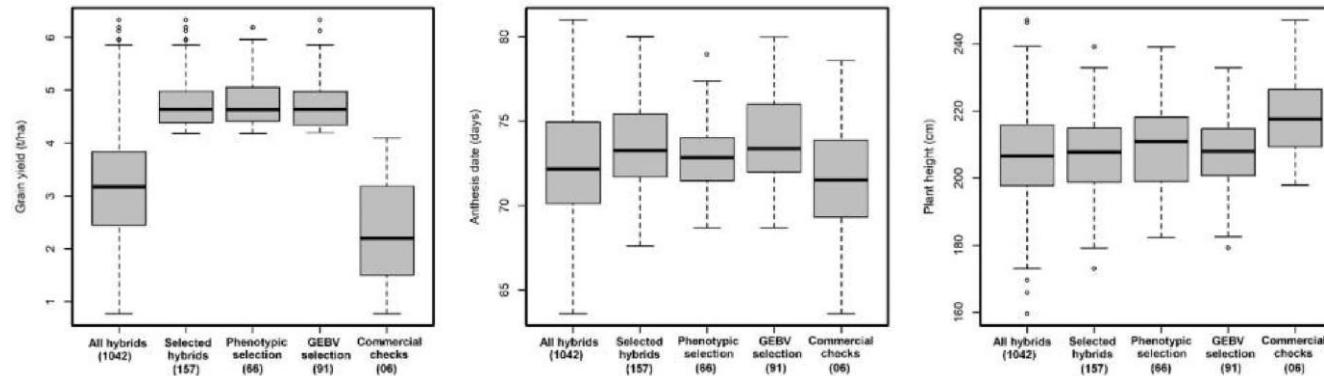
Anthony T. Slater , Noel O. I. Cogan & John W. Forster

Molecular Breeding 32, 299–310 (2013) | [Cite this article](#)

Model 1 conventional selection to G2 followed by disease screening				Model 2 conventional selection and G2 marker application				Model 4 G0 MAS then conventional screening without increasing the population			
	Intense	Moderate	Mild		Intense	Moderate	Mild		Intense	Moderate	Mild
<i>G0 cost</i>				<i>G0 cost</i>				<i>G0 cost</i>			
No. of seedlings ^a	100,000	20,000	6,667	No. of seedlings	100,000	20,000	6,667	No. of seedlings	100,000	20,000	6,667
Seedling tuber production	\$300,000	\$60,000	\$20,001	Seedling tuber production	\$300,000	\$60,000	\$20,001	Seedling tuber production	\$300,000	\$60,000	\$20,001
<i>G1 cost</i>				<i>G1 cost</i>				<i>G1 marker application</i>			
No. of seedlings	100,000	20,000	6,667	No. of seedlings	100,000	20,000	6,667	No. of seedlings	100,000	20,000	6,667
Plant spacing (m)	0.6	0.6	0.6	Plant spacing (m)	0.6	0.6	0.6	96 well plate samples	90	90	90
Row spacing (m)	0.82	0.82	0.82	Row spacing (m)	0.82	0.82	0.82	No. of plates	1,111	222	74
Paddock length (m)	100	100	100	Paddock length (m)	100	100	100	Cost per plate	736	736	736
Total area (ha)	4.92	0.98	0.33	Total area (ha)	4.92	0.98	0.33	Total marker cost	\$817,778	\$163,556	\$54,521
G1 cost ^b	\$164,731	\$32,946	\$10,983	G1 cost ^b	\$164,731	\$32,946	\$10,983	Simplex segregation	46.4 %	46.4 %	46.4 %
G1 selection rate	2 %	10 %	30 %	G1 selection rate	2 %	10 %	30 %	<i>G1 seedlings</i>			
<i>G2 cost</i>				<i>G2 cost</i>				No. G1 seedlings	46,400	9,280	3,094
No. of genotypes	2,000	2,000	2,000	No. of genotypes	2,000	2,000	2,000	Plant spacing (m)	0.6	0.6	0.6
Genotype spacing (m)	5	5	5	Genotype spacing (m)	5	5	5	Row spacing (m)	0.82	0.82	0.82
Row spacing (m)	2.2	2.2	2.2	Row spacing (m)	2.2	2.2	2.2	Paddock length (m)	100	100	100
Paddock length (m)	100	100	100	Paddock length (m)	100	100	100	Total area (ha)	2.28	0.46	0.15
No. of rows	100	100	100	No. of rows	100	100	100	G1 cost ^a	\$76,435	\$15,287	\$5,097
Total area (ha)	2.2	2.2	2.2	Total area (ha)	2.2	2.2	2.2	G1 selection rate	2 %	10 %	30 %
G2 cost ^c	\$119,552	\$119,552	\$119,552	G2 cost ^c	\$119,552	\$119,552	\$119,552	<i>G2 cost</i>			
G2 selection rate	10 %	10 %	10 %	G2 selection rate	10 %	10 %	10 %	No. of genotypes	928	928	928
<i>G3 disease screening for PCN resistance</i>				<i>G2 marker application</i>				Plant spacing (m)	5	5	5
No. of genotypes	200	200	200	No. of genotypes	2,000	2,000	2,000	Row spacing (m)	2.2	2.2	2.2
Cost per cultivar	\$219.20	\$219.20	\$219.20	No. samples per 96 well plate	90	90	90	Paddock length (m)	100	100	100
PCN screening trial cost	\$43,840	\$43,840	\$43,840	No. of plates	22.2	22.2	22.2	Total area (ha)	1.02	1.02	1.02
				Cost per plate	736	736	736	G2 cost ^b	\$55,472	\$55,472	\$55,472
				Total marker cost	\$16,356	\$16,356	\$16,356	<i>Total cost</i>	\$1,249,685	\$294,315	\$135,091
<i>Total cost</i>	\$633,040	\$257,248	\$194,618	<i>Total cost</i>	\$605,556	\$229,764	\$167,133				

- Phenotypic screen for disease.
- 4 generations (G0 - G3).
- Cost = AUD 195k - 633k
- MAS at G2
- 3 generations.
- Cost = AUD 167k - 606k
- MAS at G0.
- 3 generations.
- Cost = AUD 135k – 1,250k

Cost comparison (an example in tropical maize breeding)



Empirical Comparison of Tropical Maize Hybrids Selected Through Genomic and Phenotypic Selections

Yoseph Beyene^{1*}, Manje Gowda¹, Michael Olsen¹, Kelly R. Robbins²,
Paulino Pérez-Rodríguez³, Gregorio Alvarado⁴, Kate Dreher⁴, Star Yanxin Gao²,
Stephen Mugo¹, Boddupalli M. Prasanna¹ and Jose Crossa⁴

TABLE 5 | Cost–benefit analysis of phenotypic selection and genomic selection in International Maize and Wheat Improvement Center's (CIMMYT's) maize breeding program in Kenya.

Methods	Cost/entry (US\$)	No. of entries	No. of reps/sites	No. of rows/sites	No. of sites	Total cost (US\$)
PS (making testcrosses)	10	1492	1	1	1	14,920
PS (stage I multi-location yield trials)	5	1492	2	2	4	119,360
GS (making testcrosses)	10	855	1	1	1	8,550
GS (phenotyping training set in stage I multi-location yield trials)	5	855	2	2	4	68,400
GS (genotyping all lines)	10	1492	1	1	1	14,920
Total cost of GS						91,870
Total cost of PS						134,280
GS:PS cost ratio						0.68

Available infrastructure/technology

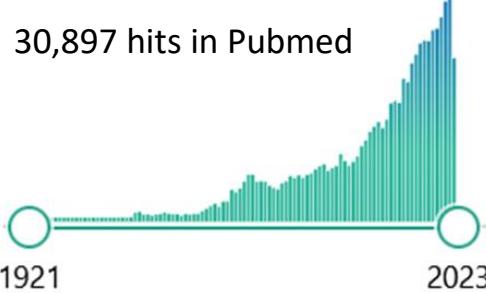
We may not always have access to everything:

- Field
- Greenhouse, growth room/chamber
- Hybrid production system (manual, CMS, ...)
- Mutagen (EMS: ethyl methanesulfonate, gamma ray, cosmic ray, ...)
- In-house genotyping facility (often outsourced)
- Computational facility (cloud computing is an alternative option)
- Laboratory facility for genetic transformation and tissue culture.
- Phenotyping platforms (enclosed system, drone, satellite, spectroscopy, ...)

Most of these can be solved by outsourcing (increased cost).

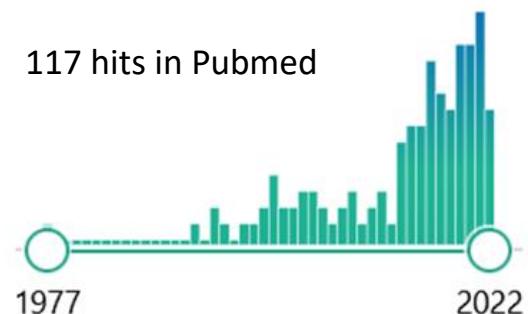
Current knowledge – two contrasting examples

Soybean (*Glycine max*)



<https://phys.org/news/2021-07-soybean-expansion-south-america-years.html>

Dulse (*Palmaria palmata*)



Genome size	1 Gb	780 Mb (???)
Chromosome	$2n = 40$	$2n/4n = 48$ (???)
Reproduction	Hermaphrodite, self pollinate	Dioecious, outcross(?)
Growth cycle	3 months	1 year (but can be harvested earlier)
Habitat	Land	Sea (rocks, kelps)
Production	Farmed worldwide	Wild harvested in East Asia, North America, Europe
Key trait	Bean yield (known genetics)	Complete biomass yield (unknown genetics)
Consumption	Edamame, soy milk, tofu, ...	Laver bread, “vegan bacon”, seasoning, ...

Current knowledge

- Crop knowledge determines what we can, and cannot, do with the crop.
- Most of the conventional and advanced breeding methods can be applied in soybean.
- However, the breeding options are limited for dulse.

Ready: DH (shortcut), clonal

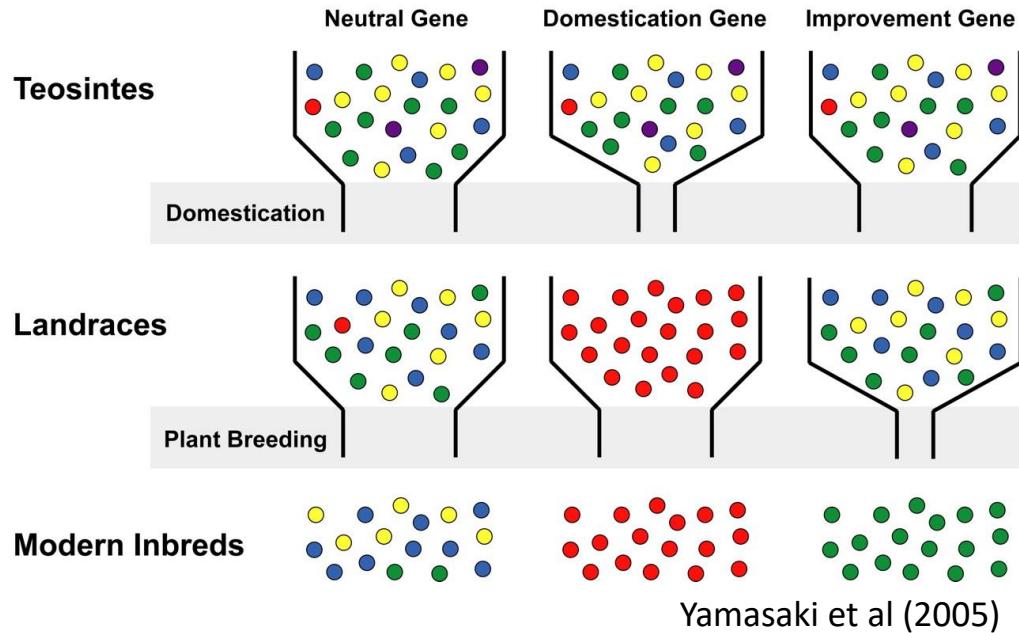
Maybe: pedigree, SSD, bulk, backcross, pop/evo, hybrid, mutation

Not today: MAS, GS, GWAS, Linkage mapping, GM, GE, Omics

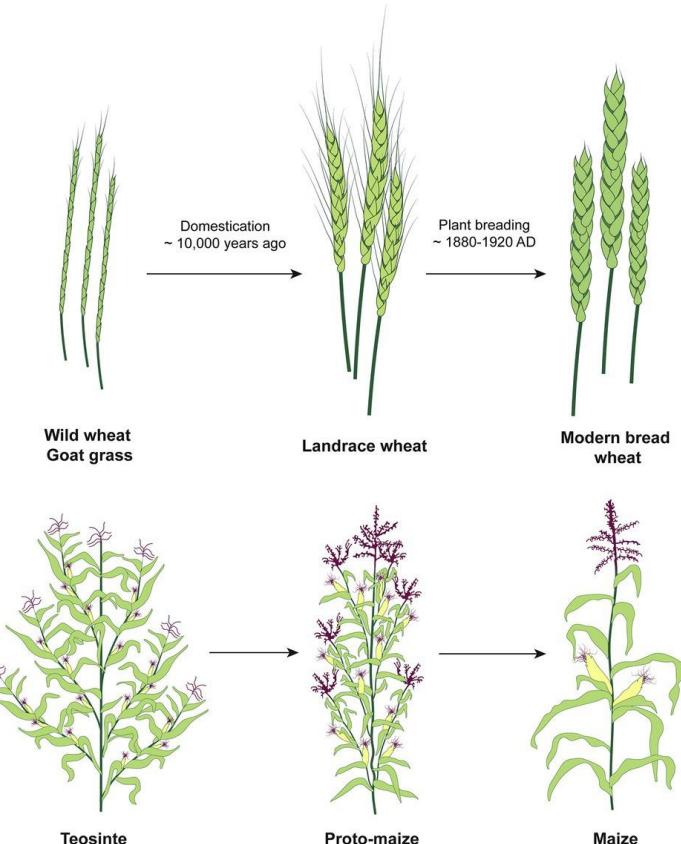
- As you may have noticed, the advanced breeding methods rely heavily on the crop knowledge.
- Soybean vs dulse are on the two extreme ends – many crops are in between.

Genetic diversity

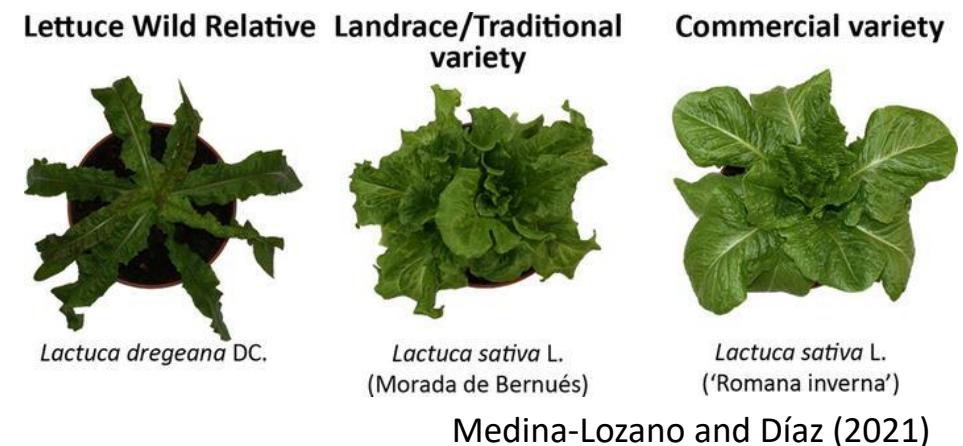
Domesticated crop generally follows: Wild > Landrace > Cultivar



Diversity is highest in wild, reduced in landrace and further reduced in modern cultivars.



Pourkheirandish et al (2020)



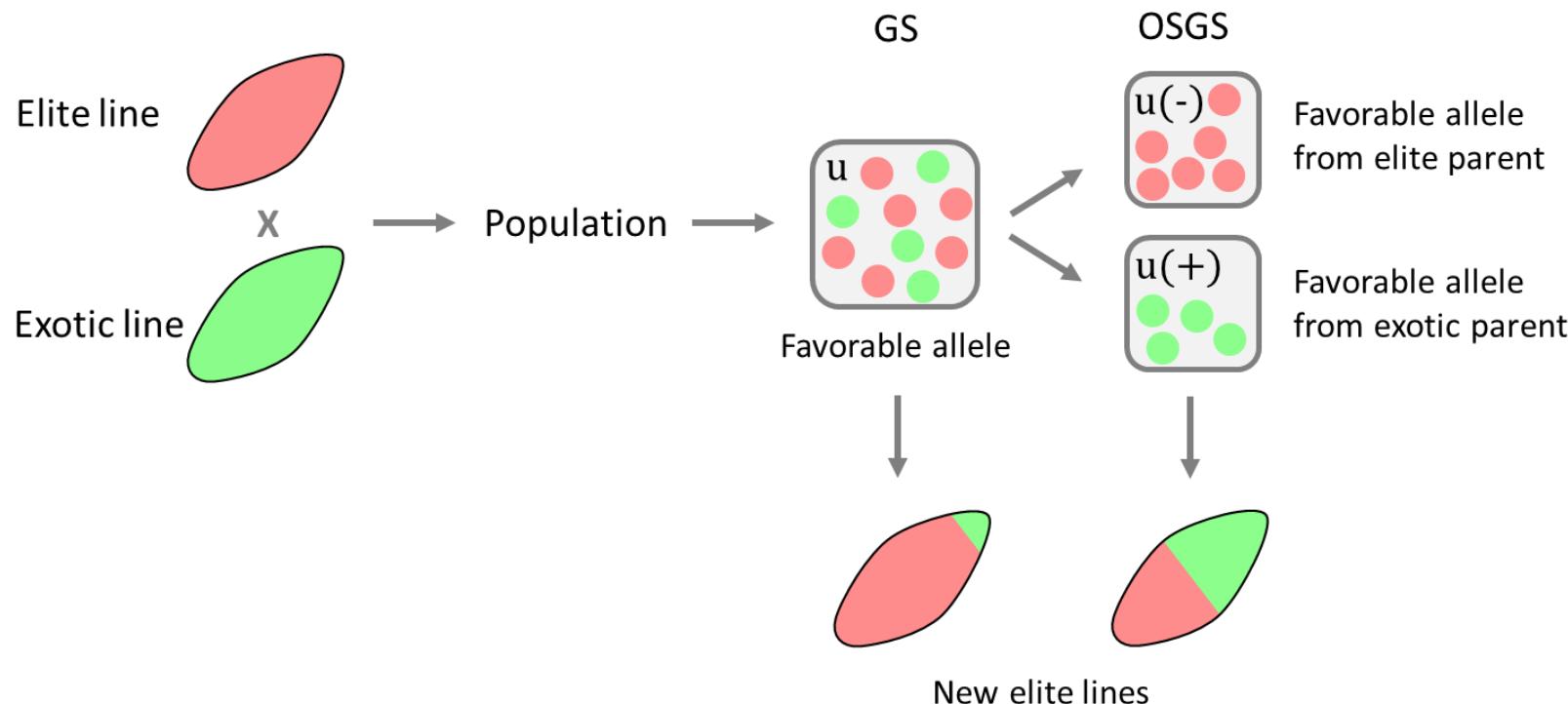
Genetic diversity

What to do when our modern breeding population is lacking in diversity?

1. Create diversity: mutation, GM, GE.
2. Shuffle diversity: backcross + MAS (QTLs identified from GWAS/Linkage mapping), OSGS.

OSGS: Origin specific genomic selection

- Allows us to identify and target selection on favorable exotic alleles from wild/landrace.



Bandwagon and hype

Bandwagons I, too, have known

Rex Bernardo 

Theoretical and Applied Genetics 129, 2323–2332 (2016) | [Cite this article](#)

5234 Accesses | 126 Citations | 35 Altmetric | [Metrics](#)

Abstract

Key message

Bandwagons come in waves. A plant breeder, just like a surfer, needs to carefully choose which waves to be on.

- The hype (excitement) phase of a bandwagon is important.
- It is the best opportunity for getting funding.
- In a way, bandwagon pushes advanced breeding methods forward.

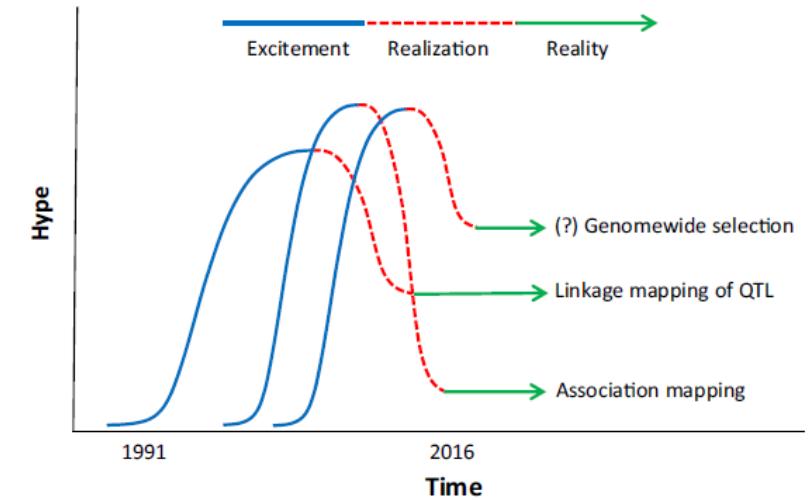
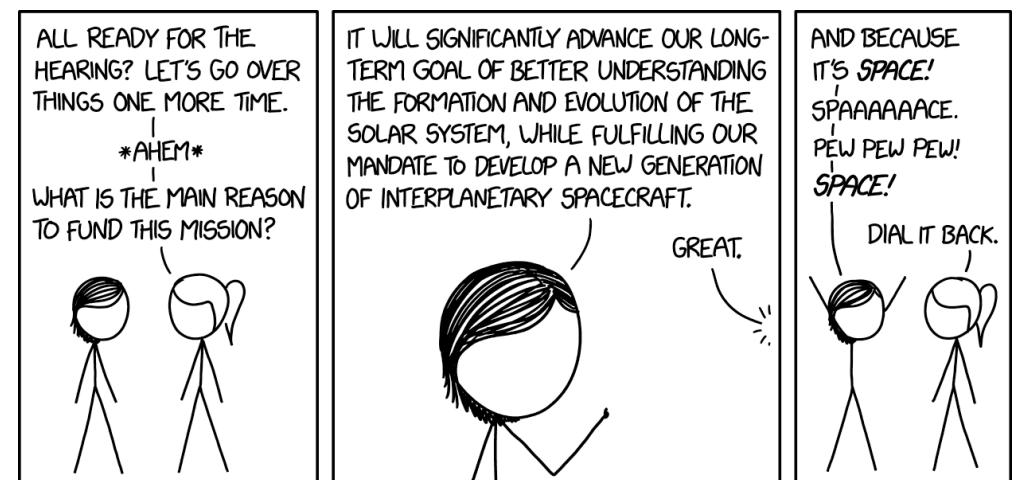


Fig. 1 Life cycle of a bandwagon, with QTL mapping, association mapping, and genomewide selection as examples. The reality level for association mapping is low, because the approach typically lacks power for detecting rare variants, which are what plant breeders most often seek. The (?) before genomewide selection indicates that the eventual level of usefulness of the procedure is still being discovered



The breeder's equation

The question of conventional vs advanced breeding methods ultimately goes back to the breeder's equation.

Recall from Lecture 5: $\Delta R = \frac{ih\sigma_g}{t}$

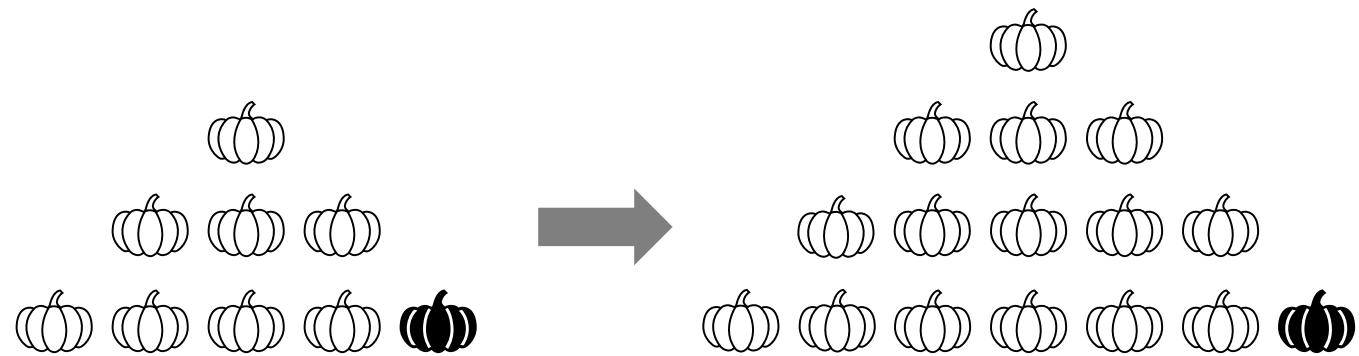
Which can be interpreted as: $Rate\ of\ genetic\ gain = \frac{selection\ intensity \times selection\ accuracy \times genetic\ variation}{time}$

How can we maximize the rate of genetic gain in our breeding program?

Selection intensity

We can increase selection intensity by:

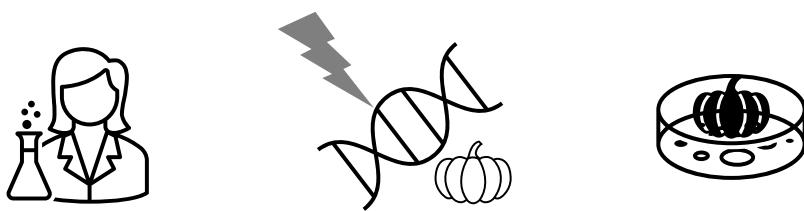
1. Increasing the population size (budget limitation!).



2. Reducing the cost of genotype evaluation (trial design, MAS, GS, phenomics).



3. Creating our target through GM/GE (maximum selection intensity).



Selection accuracy

We can increase selection accuracy by:

1. Increasing the trial locations (more trials = better estimate of breeding value, budget limitation!).
2. Improving breeding value estimate using MAS/GS.
3. Selecting on correlated traits with higher h^2 .
4. Using phenomics (model for crop development).
5. Digitization and better management of data (fits into computational pipeline).
6. Digital twinning – real time modelling of crop using actual data and simulation.
7. Precision, customer-oriented breeding (improves directly on the desired traits, can be done with MAS and GE).

Genetic variation

We can increase/maintain genetic variation by:

1. Creating/shuffling genetic variation (mutation, GM, GE, MAS, OSGS, as discussed earlier).
2. Using recurrent selection (stacking of favorable alleles over time).
3. Using optimum contribution selection (OCS) – controls the loss of genetic variation due to selection.

Time

We can reduce generation time by:

1. Using alternative method for genotype evaluation (e.g. MAS/GS, potato example).
2. Using rapid cycling (shortened version of pedigree breeding which takes many years for a single cycle).
3. Hastening crop development through SSD/speed breeding.

In reality...

Conventional and advanced breeding methods are not mutually exclusive!



Conventional breeding + MAS/GS

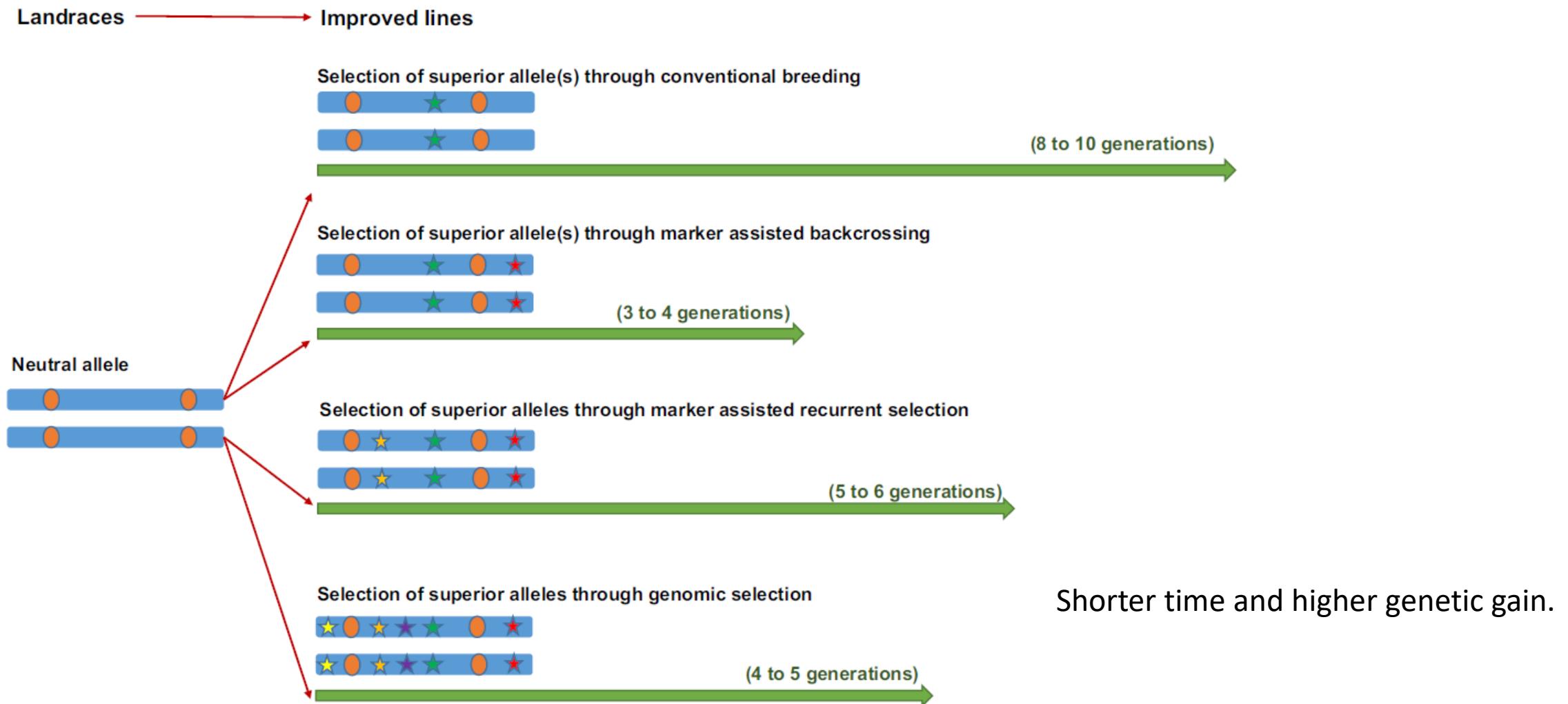


Figure 4 Representation of different approaches of transferring elite alleles through genomics-assisted breeding with approximate timescale. The oval shapes represent the neutral alleles and stars represent the superior alleles of interest in plant-breeding programs. The timescale shows the approximate time required by breeding program to achieve introgression of superior allele from landraces to the improved lines.

Evolution of breeding program

Fifty years of a public cassava breeding program:
evolution of breeding objectives, methods, and decision-making processes

Hernán Ceballos , Clair Hershey, Carlos Iglesias & Xiaofei Zhang

Theoretical and Applied Genetics 134, 2335–2353 (2021) | [Cite this article](#)

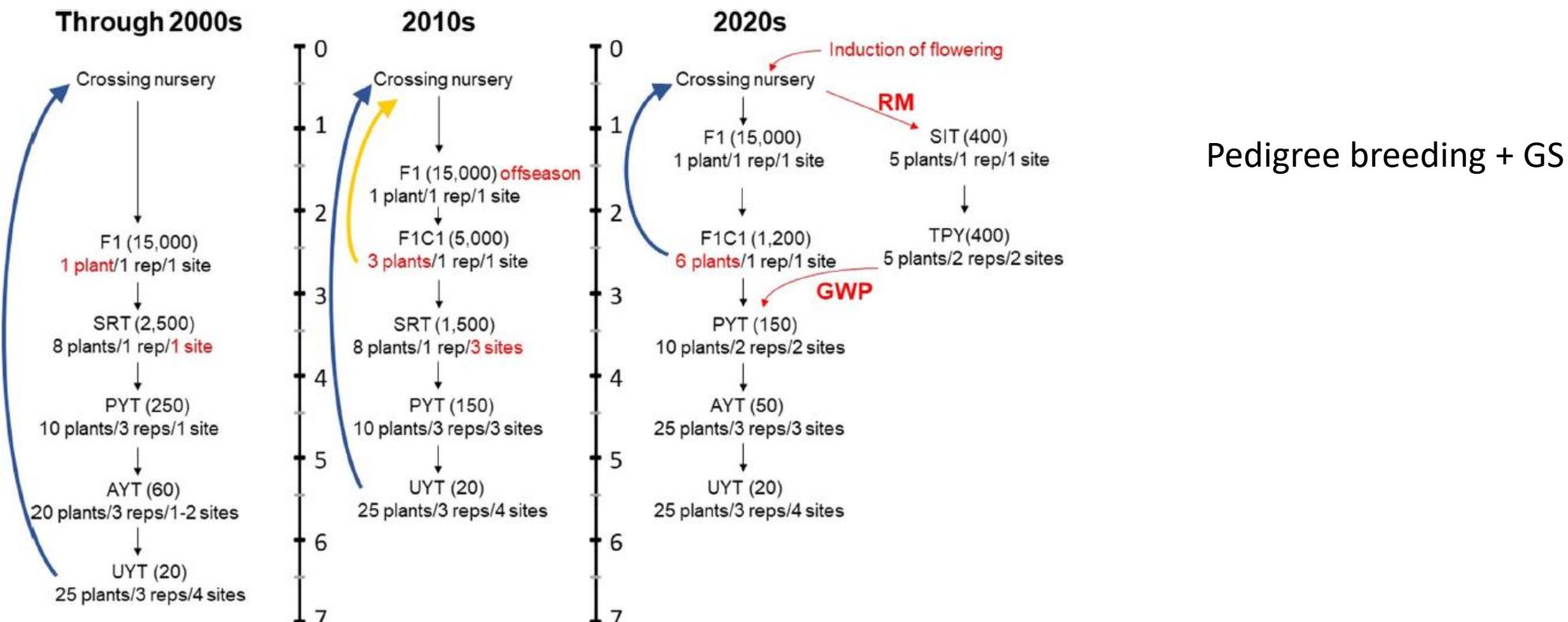


Fig. 1 The evolution of breeding schemes at CIAT. F1: seedling nursery; F1C1: cloned seedling nursery; SRT: single row trial; PYT, AYT and UYT: preliminary, advanced and uniform yield trials, respectively; GWP: genome-wide prediction; SIT: seed increase trial; TPY:

training population yield trials; RM: rapid multiplication. Upward arrows indicate the duration of each breeding cycle (the orange arrow is for the rapid-cycling for high-carotenoids breeding)

Example of MAS + backcross in maize

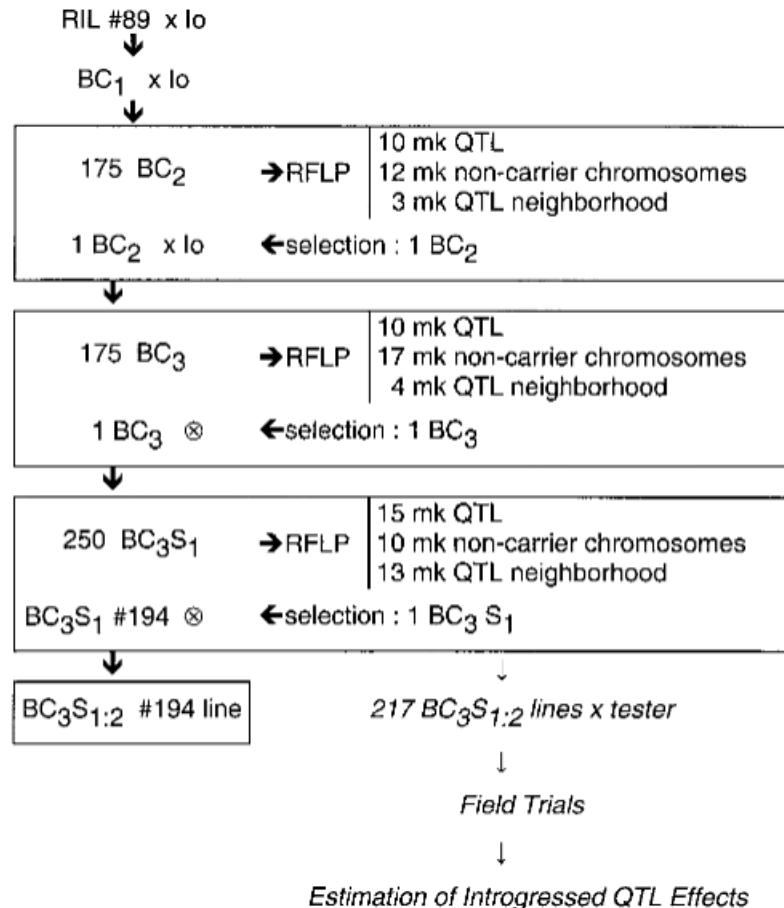
Marker-Assisted Introgression of Favorable Alleles at Quantitative Trait Loci Between Maize Elite Lines

Agnès Bouchez,¹ Frédéric Hospital, Mathilde Causse,² André Gallais and Alain Charcosset

Station de Génétique Végétale, INRA-UPS-INAPG, Ferme du Moulin, 91190 Gif-sur-Yvette, France

Manuscript received January 7, 2002

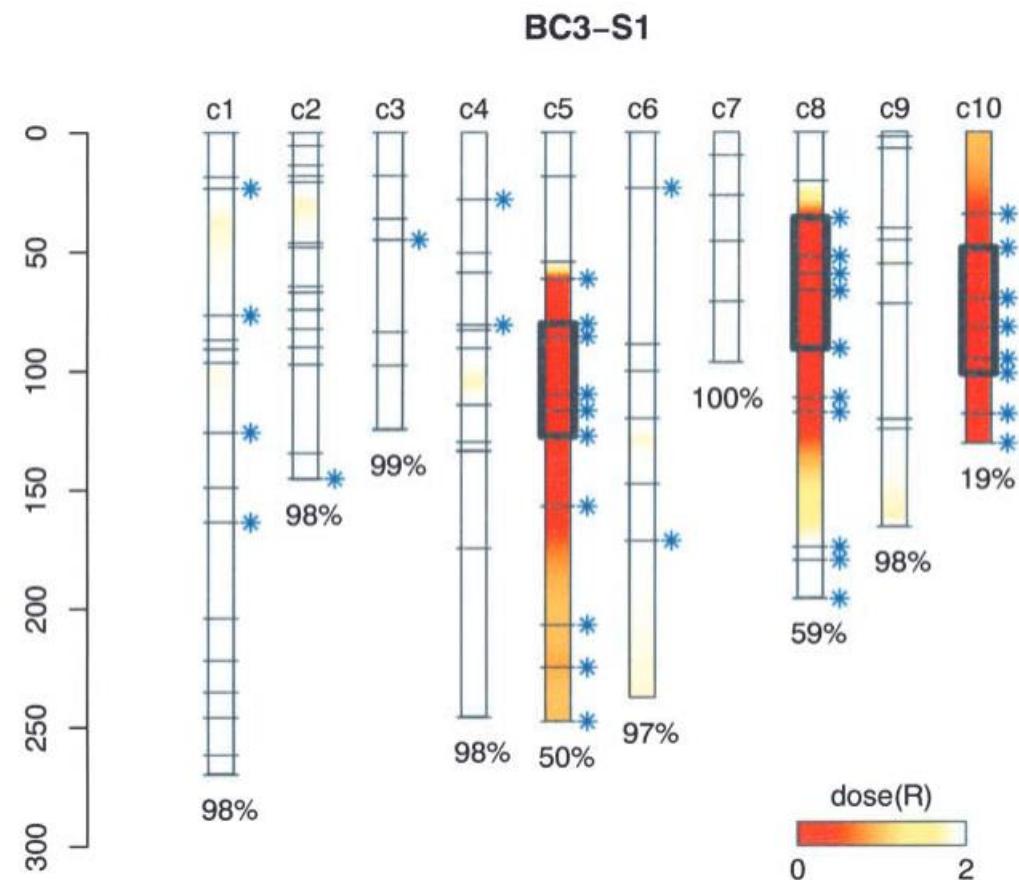
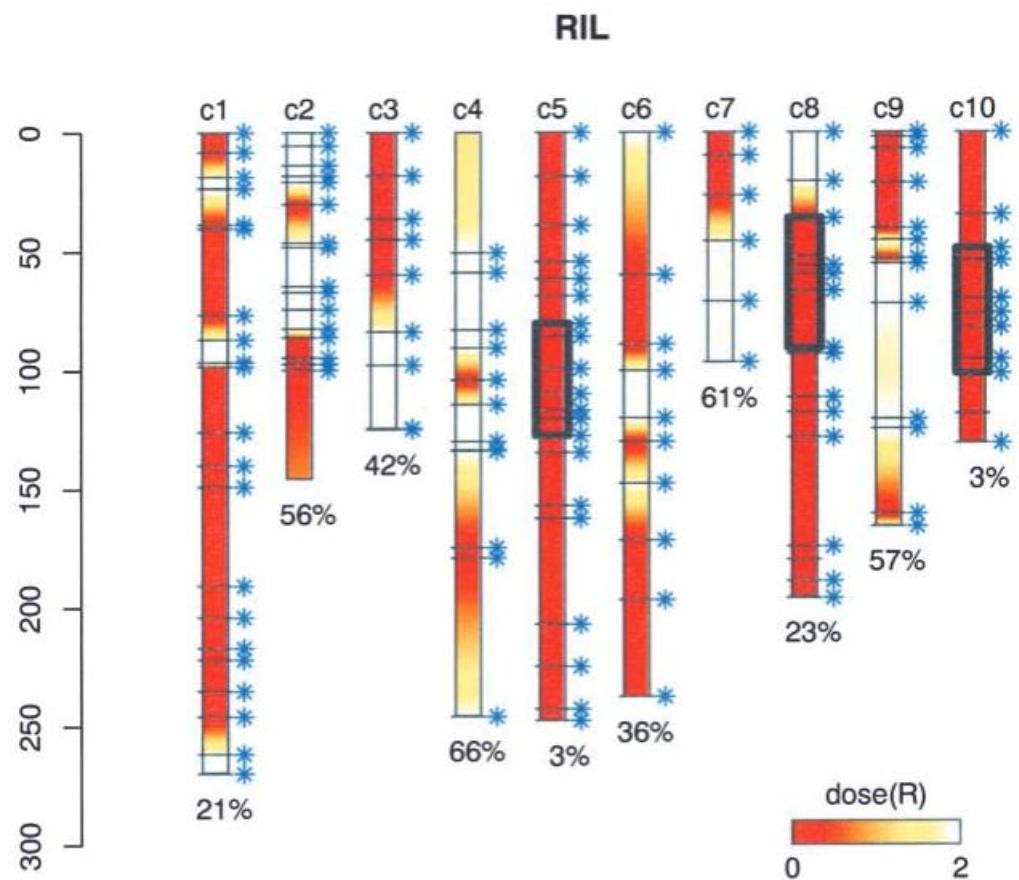
Accepted for publication September 19, 2002



- MAS + backcross to introgress 3 QTLs for silking date (SD), grain moisture (GM) and dry grain yield (DGY).
- Foreground vs background selection.
- Goal 1: transfer the QTLs from donor line to recipient line.
- Goal 2: recover the recipient genome elsewhere.

FIGURE 2.—Marker-assisted backcross introgression scheme.

Example of MAS + backcross in maize



RIL: donor line (QTLs highlighted in blue rectangles).

BC3-S1: result of MAS + backcross.

Example of MAS + backcross in maize

TABLE 5
Summary of introgression effects

	Results	Trait			
		SD (days)	GM (%)	DGY (qx/ha)	Economic index (qx/ha)
Predicted from QTL effects (RIL)	Table 1	-1.6	-1.6	+12.4	+16.4
Estimated from QTL effects (BC_3S_1)	Table 4	-3.6	-1.9	-7.1	-2.3
Observed phenotypic mean for $BC_3S_{1.2}$ no. 194	Table 3	-4.1	-1.9	-12.9	-8.1

Summary of predicted, estimated, and observed introgression effects for SD (silking date), GM (grain moisture), DGY (dry grain yield), and economic index (DGY - 2.5 GM, see QTL DETECTION AND INTROGRESSION). Overall additive effects of F2 alleles were obtained from reference effects in Table 1 and from effects for groups of all trials in Table 4.

The ending did not turn out as expected...

Which trait has gone wrong?

Example of GM + hybrid in papaya

‘UH Rainbow’ Papaya

A High-Quality Hybrid with Genetically Engineered Disease Resistance

Richard M. Manshardt, Department of Horticulture

‘UH Rainbow’ papaya is a new hybrid variety developed by the College of Tropical Agriculture and Human Resources of the University of Hawaii at Manoa. ‘UH Rainbow’ combines the superior quality typical of Hawaii’s “solo” papayas with excellent resistance to a devastating plant virus disease—papaya ringspot virus (PRSV). This combination of traits was accomplished through genetic engineering, one of the latest advances in agricultural biotechnology (see page 2, over).

‘UH Rainbow’ is an F₁ (first-generation) hybrid produced by crossing Hawaii’s standard export variety, ‘Kapoho’, with the first genetically engineered papaya with resistance to PRSV, ‘UH SunUp’. The resulting hybrid is an excellent source of vitamins A and C and is highly productive.



CTAHR University of Hawaii (1999)

Example of GM + hybrid in papaya



Figure 2. PRSV infected papaya fields in 1994.

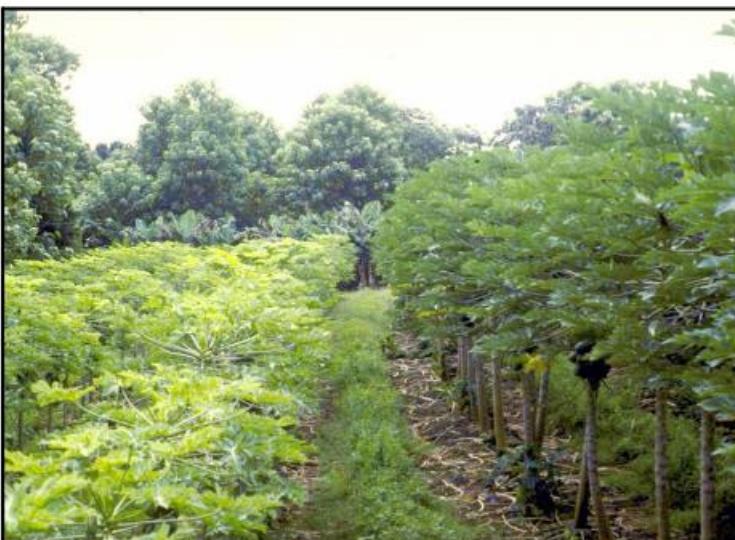


Figure 3. Transgenic papaya test field in Puna. Yellow plants are nontransgenic papaya; plants on right are transgenic 'Rainbow' papaya. Photo by S. Ferreira.



Figure 4. Aerial view of transgenic papaya test field in Puna, showing block of healthy transgenic 'Rainbow' surrounded by severely infected nontransgenic 'Sunrise' papaya. Photo by S. Ferreira.

Gonsalves (2004)

Rainbow contains resistance to papaya ring spot virus (PRSV).

First GE crop

NEWS | 14 December 2021

GABA-enriched tomato is first CRISPR-edited food to enter market

Sanatech Seed's Sicilian Rouge CRISPR-edited 'health-promoting' tomatoes reach consumers and may open the market to more genome-edited fruit, vegetables and even fish.

Emily Waltz



Genome-edited food made with CRISPR–Cas9 technology is being sold on the open market for the first time. Since September, the Sicilian Rouge tomatoes, which are genetically edited to contain high amounts of γ -aminobutyric acid (GABA), have been sold direct to consumers

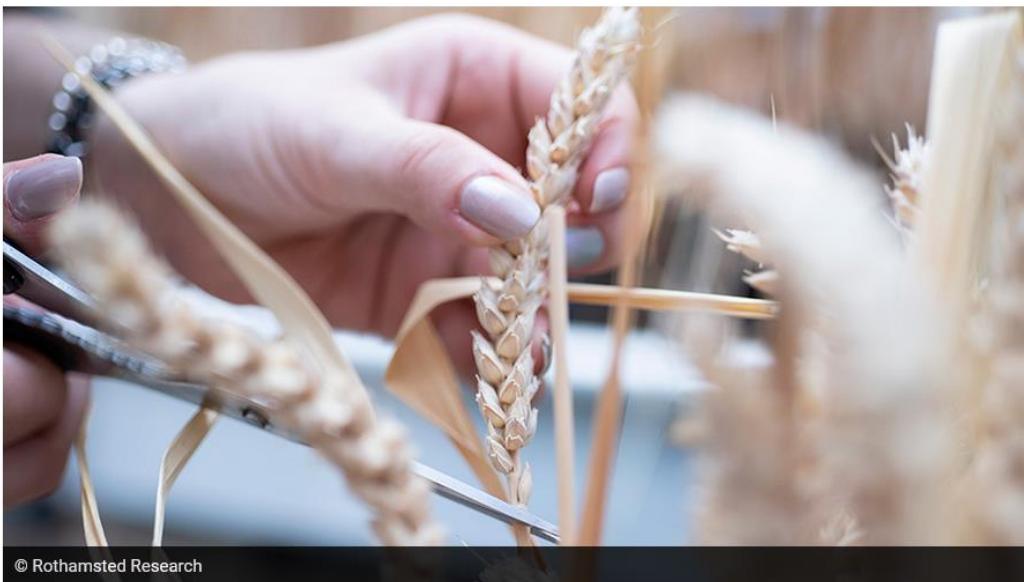


Sanatech

GE is fairly new - may take few years before we start seeing routine use in crop breeding.

A closer example: GE in Cadenza wheat

How growers could benefit from pioneering gene-edited wheat



© Rothamsted Research

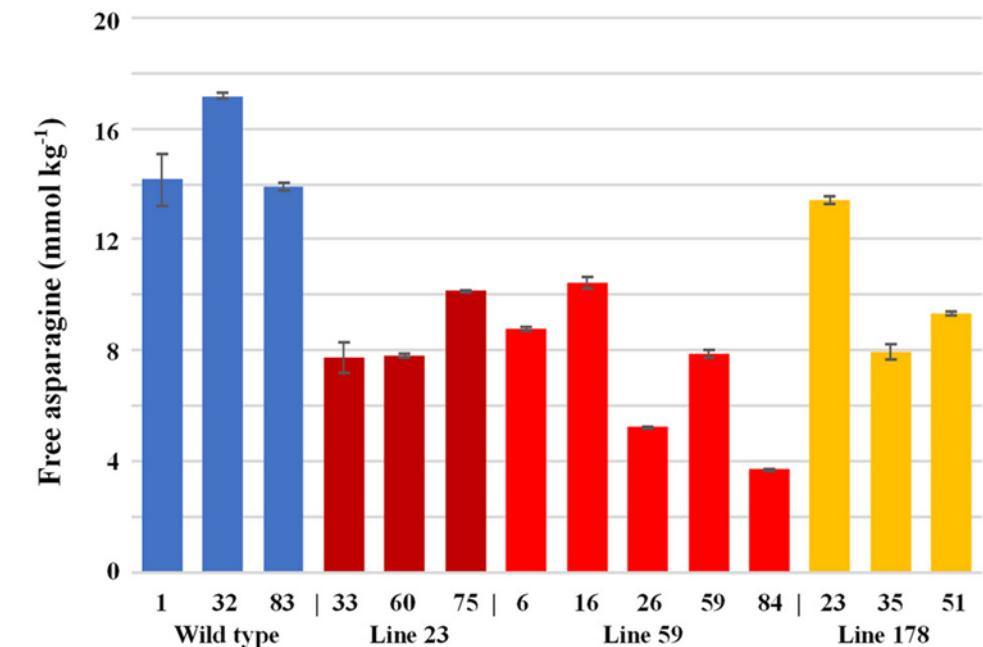
Wheat bred for low levels of a key amino acid could prove vital for food manufacturers meeting future limits on acrylamide, especially in breakfast cereals and biscuits.

<https://www.fwi.co.uk/arable/wheat/how-growers-could-benefit-from-pioneering-gene-edited-wheat>

Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene *TaASN2*

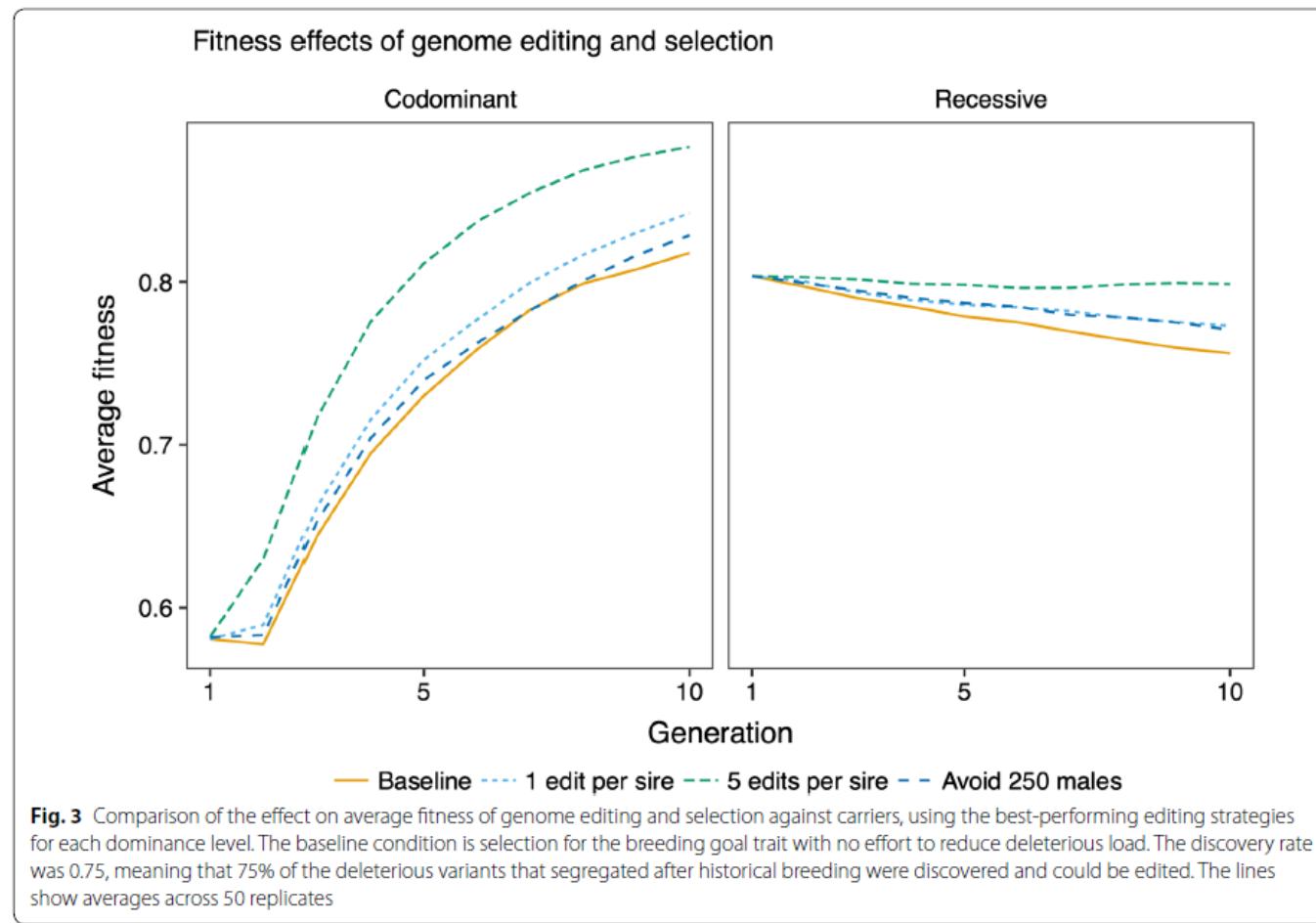
Sarah Raffan, Caroline Sparks, Alison Huttly, Lucy Hyde, Damiano Martignago, Andrew Mead, Steven J. Hanley, Paul A. Wilkinson, Gary Barker, Keith J. Edwards, Tanya Y. Curtis, Sarah Usher, Ondrej Kosik, Nigel G. Halford ... See fewer authors ^

First published: 26 February 2021 | <https://doi.org/10.1111/pbi.13573> | Citations: 16



Simulation study on GE in removing deleterious alleles.

Compare the effectiveness in using GE vs selection to remove deleterious alleles.



Johnsson et al (2019)

Knowledge exchange between plant and animal breeding

Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery

John M Hickey, Tinashe Chiurugwi, Ian Mackay, Wayne Powell & Implementing Genomic Selection in CGIAR Breeding Programs Workshop Participants

Nature Genetics 49, 1297–1303 (2017) | Cite this article

8175 Accesses | 166 Citations | 90 Altmetric | Metrics

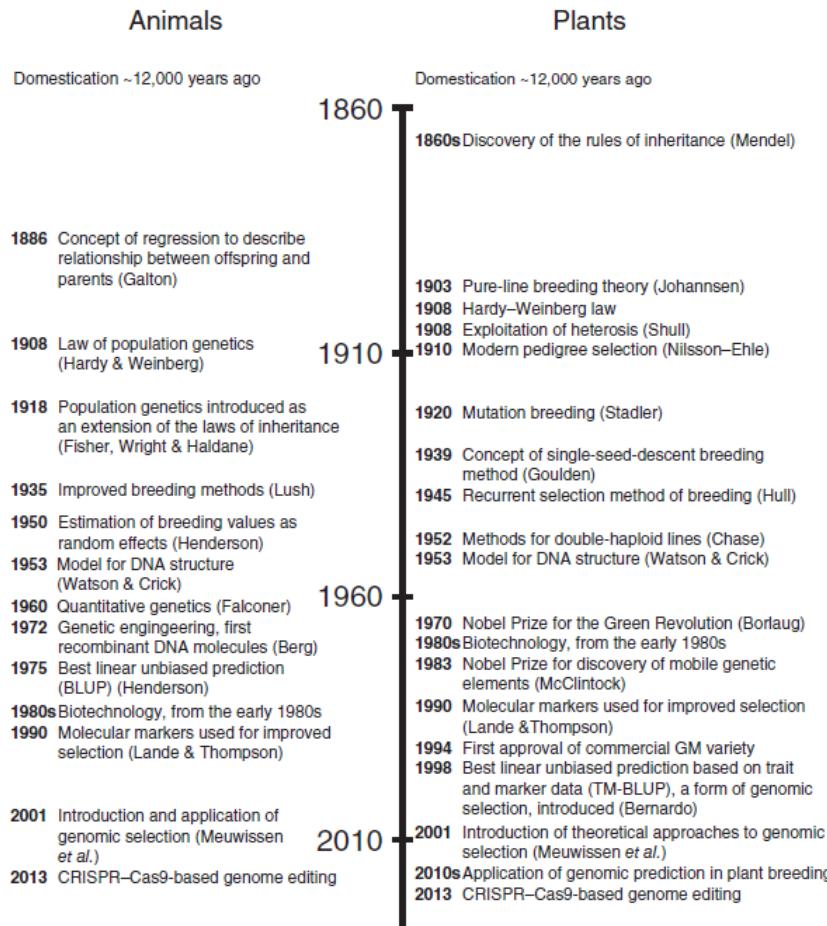


Figure 1 Some key milestones of selective animal and plant breeding.

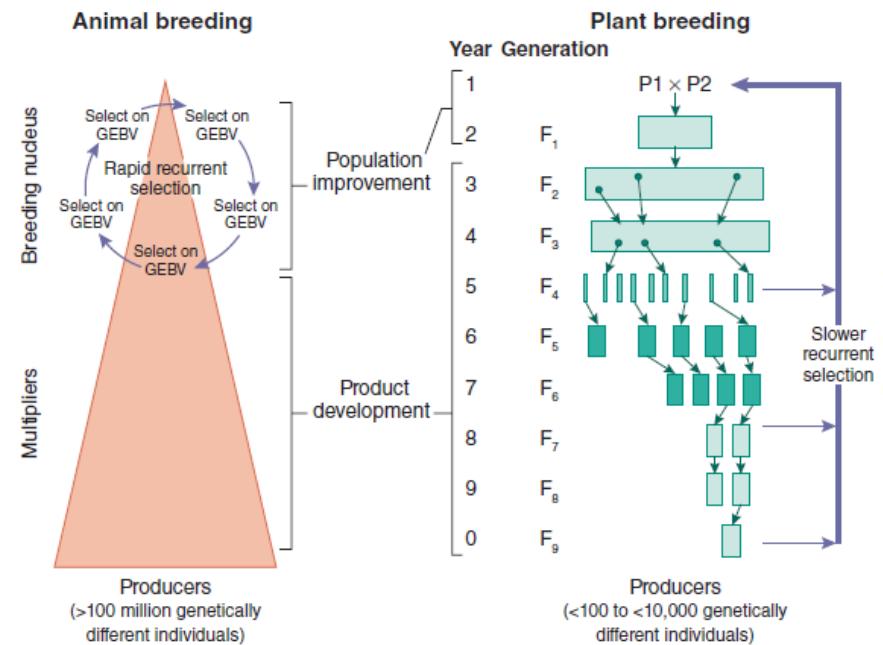


Figure 2 Comparison of animal and plant breeding approaches. In some areas, plant and animal breeders have adopted similar approaches. All breeding programs can be thought of as having two basic components: (i) a recurrent selection component that seeks to increase the frequency of favorable alleles in a population, which in turn increases the mean genetic merit of the population, and (ii) a product development component that seeks to extract genotypes from this improved population for a farmer to produce. A major difference between animal and plant breeding has been that the former places greater emphasis on population improvement, with product development consisting primarily of multiplication of stock that is not recycled into the breeding nucleus, whereas in the latter greater emphasis is placed on selection of an improved product in the form of a recognizable plant variety, which is often also the source of parents for the next breeding cycle. That is, while varieties (often protected by plant breeders' rights) are the focus in plant breeding, this concept is almost irrelevant in animals. GEBV, genomic estimated breeding value.

Future of plant breeding is not limited to conventional and advanced breeding, but also on knowledge exchange between plant and animal breeding, i.e. what can we learn from animal breeders, and what can we teach them, to better improve breeding.

Paper discussion

Annual Review of Genetics

On the Road to Breeding 4.0: Unraveling the Good, the Bad, and the Boring of Crop Quantitative Genomics

Jason G. Wallace,¹ Eli Rodgers-Melnick,²
and Edward S. Buckler^{3,4}

¹Department of Crop and Soil Sciences, The University of Georgia, Athens,
Georgia 30602, USA; email: jason.wallace@uga.edu

²Corteva Agriscience, DowDuPont, Johnston, Iowa 50131, USA

³United States Department of Agriculture, Agricultural Research Service,
Ithaca, New York 14853, USA

⁴Institute for Genomic Diversity, Cornell University, Ithaca, New York 14853, USA

Here is a review paper on plant breeding. We will not have the time to read everything, so we will focus on the following:

Abstract (page 1)

Figure 1 – 5

Summary (page 15)

Please summarize your understanding of each figure.

Wallace et al (2018) – Figure 1

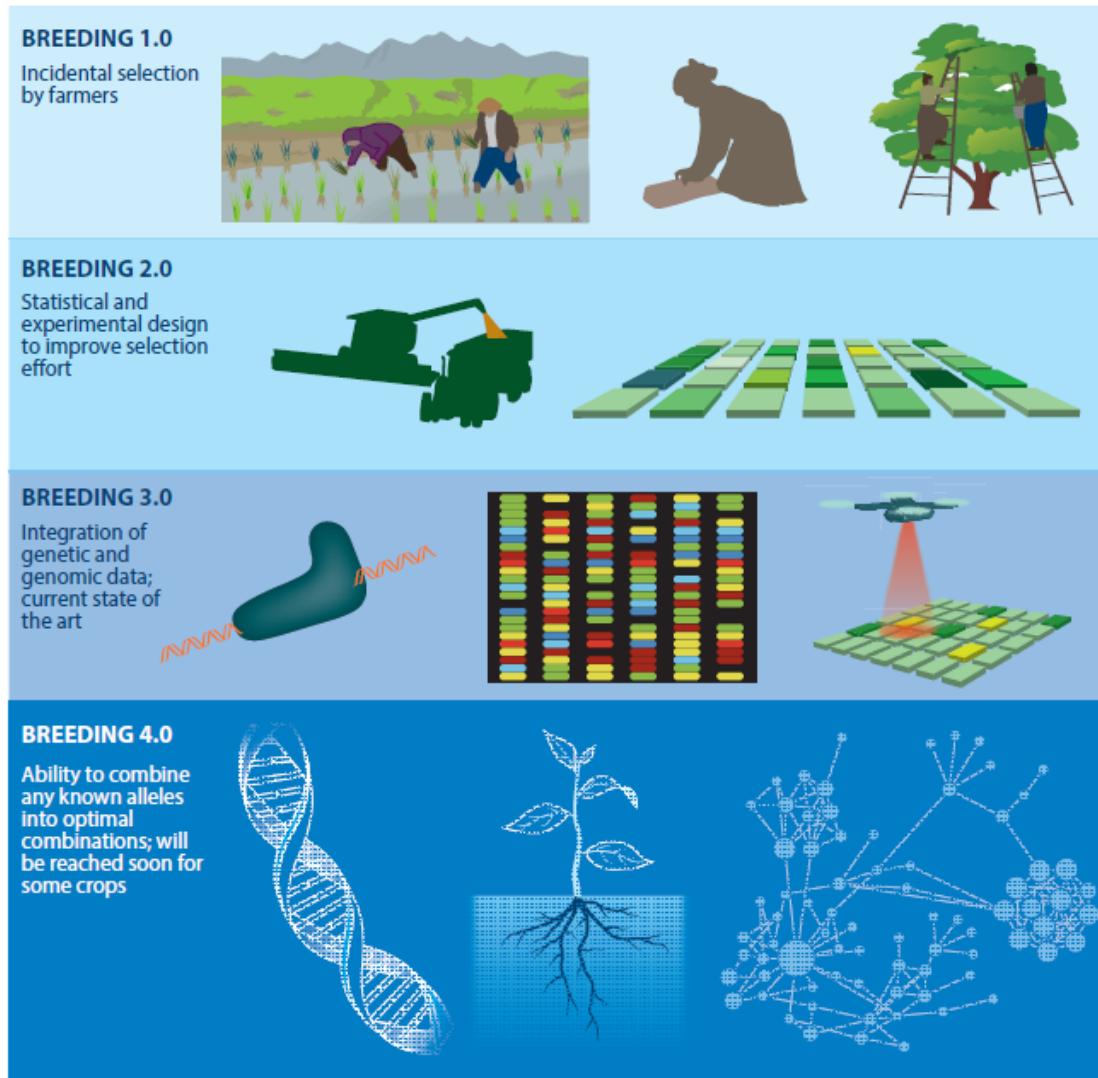


Figure 1

Four stages of plant breeding. Breeding efforts can be divided into four existing or near-future stages based on available methodology. Breeding 1.0 is mostly incidental selection by farmers. Breeding 2.0 involves using statistics and experimental design to improve selection efforts. Breeding 3.0 includes genetic and genomic data and is the current state of the art. Breeding 4.0 will probably arrive soon (at least for some crops) and is marked by the genome-wide ability to combine any known alleles into desirable combinations.

Wallace et al (2018) – Figure 2

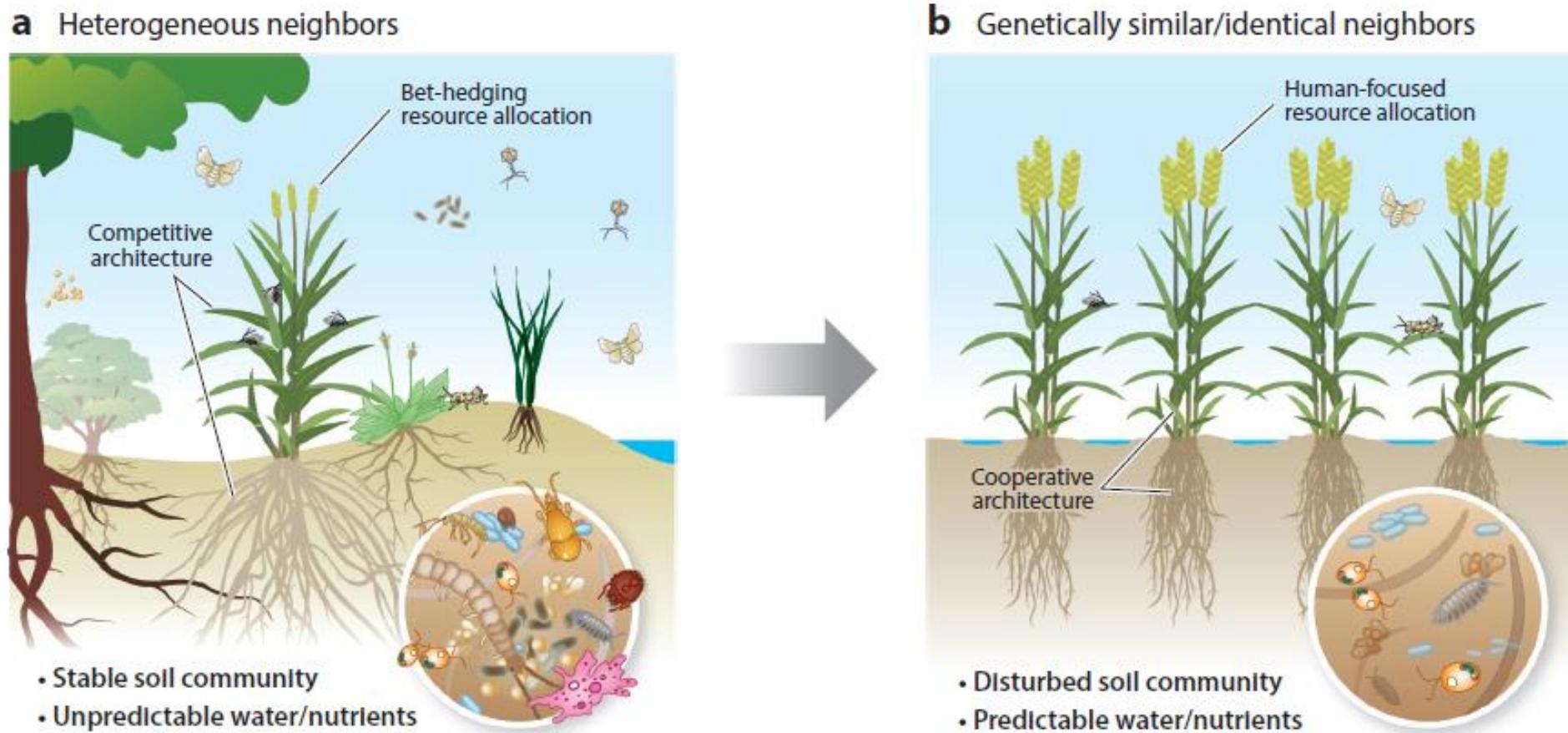


Figure 2

Plant adaptive context. (a) The environment in which crop wild ancestors evolved is very different from (b) the environment in modern agriculture. Some of the changes involved are relatively well studied, such as repartitioning resources from vegetative tissue to grain and fruit. Others, such as how plants evolve to live in a perpetually disturbed soil ecosystem, are only beginning to be understood.

Wallace et al (2018) – Figure 3

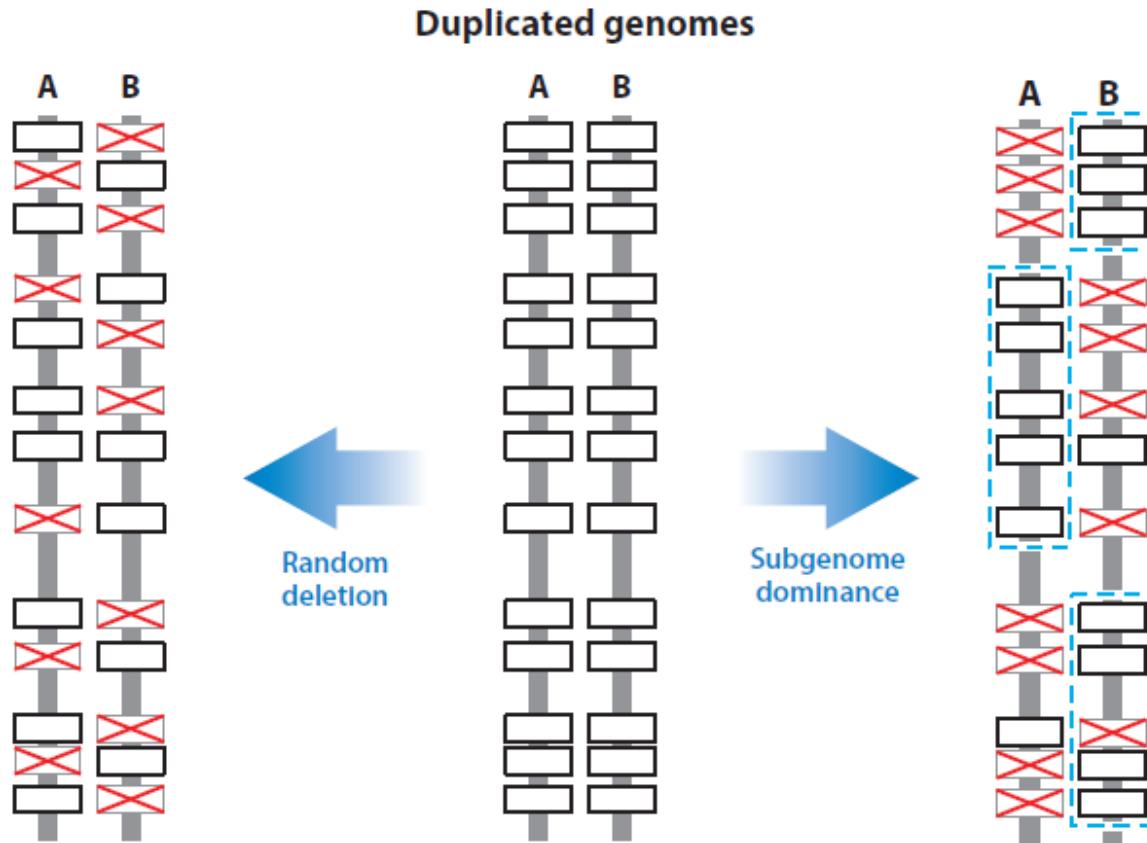


Figure 3

Subgenome dominance. Immediately after a polyploidy event, the genome contains two complete sets of genes (*A* and *B*, center). Over time, many genes are reduced to a single copy owing to mutations and deletions (red Xs). Although one would expect this process to be random (left), most genomes show evidence of genes that were preferentially retained or lost in blocks (right). The dominant subgenome consists of the blocks that are preferentially retained (blue boxes). Note that which subgenome a block belongs to depends only on how well it is retained, not on where it came from. In other words, a species with both *A* and *B* genomes due to polyploidy can have a dominant subgenome consisting of a mix of *A* and *B* segments, instead of *A* being completely dominant to *B* or vice versa.

Wallace et al (2018) – Figure 4

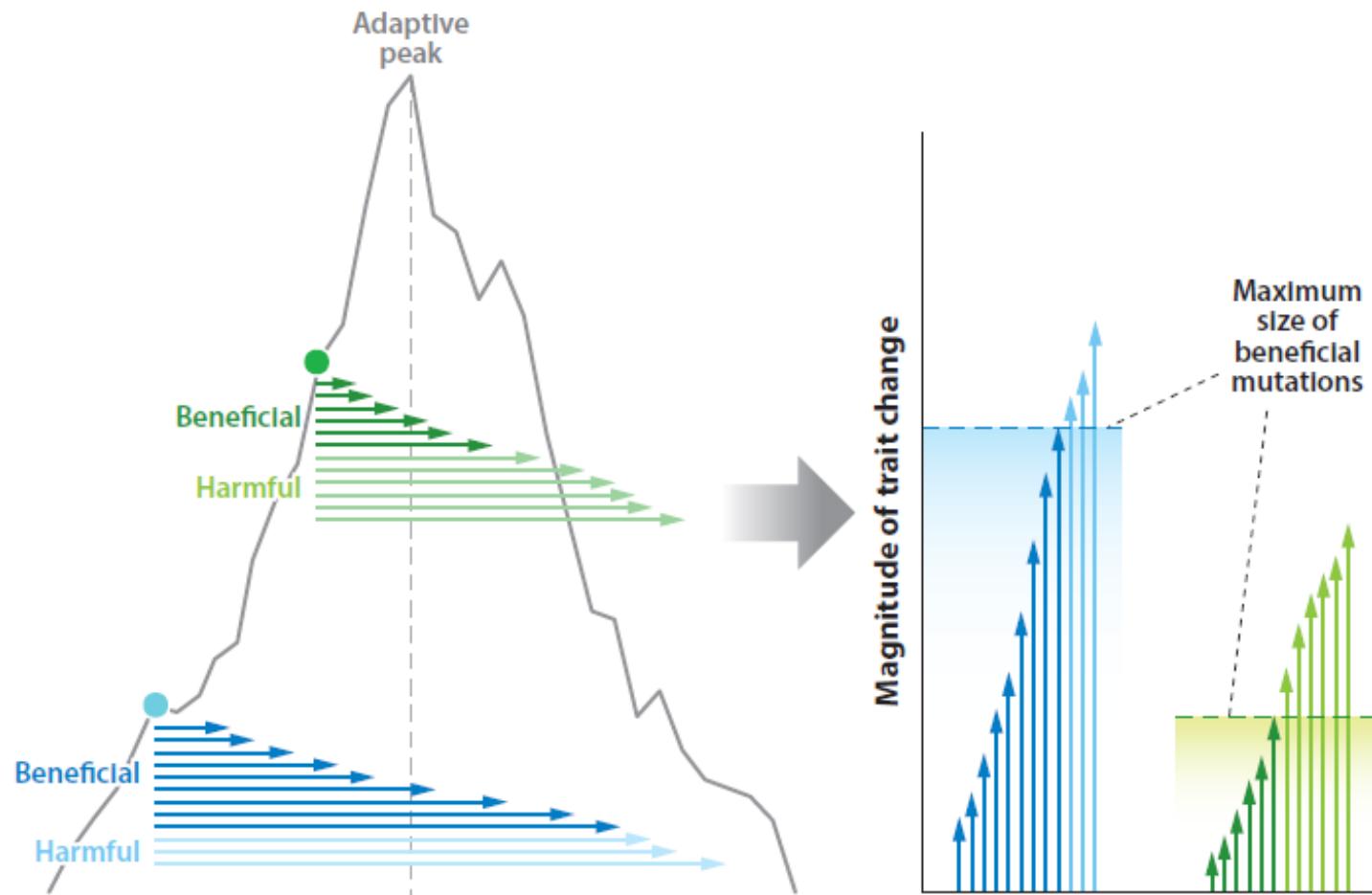


Figure 4

The Fisher-Orr model. The Fisher-Orr model states that the closer an organism is to an adaptive peak, the more likely a large mutation is to be harmful. The blue individual is far from the peak, so many different mutations result in a net benefit. By contrast, the green individual is close enough to the peak that even moderately sized changes can overshoot the ideal and leave its fitness lower than before.

Wallace et al (2018) – Figure 5

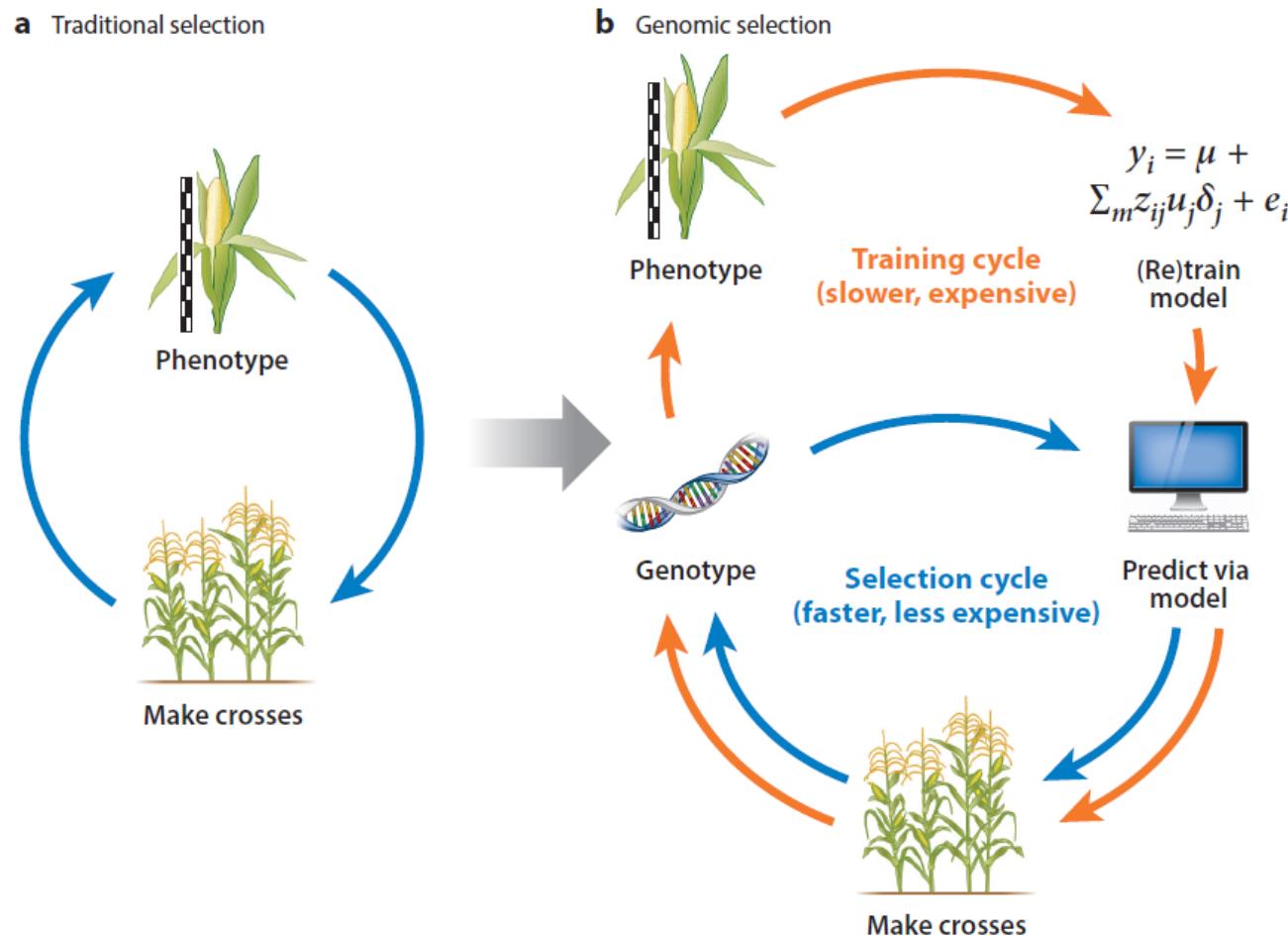


Figure 5

Genomic selection. (a) Traditional selection involves cycles of making crosses to get new genetic combinations, evaluating the new varieties by phenotyping, and using those evaluations to select the next generation of parents. (The point at which a variety is spun off into production is not shown.) (b) Genomic selection adds several layers to this scheme and splits the process into training and selection cycles. When training, the breeder must go around the entire outside loop: making crosses, getting genotype and phenotype information from them, building a mathematical model, and finally using that model to select future parents. After that, the breeder can skip the training portion and go directly from new material to genotype to choose parents based on the model, with no need to phenotype. Although the model must be retrained every few generations, breeders generally save significant time and money by skipping the (usually laborious and expensive) phenotyping step two or three times.

A Two-Part Strategy for Using Genomic Selection to Develop Inbred Lines

R. Chris Gaynor, Gregor Gorjanc, Alison R. Bentley, Eric S. Ober,
Phil Howell, Robert Jackson, Ian J. Mackay, John M. Hickey*

Gaynor et al (2017) – Figure 1

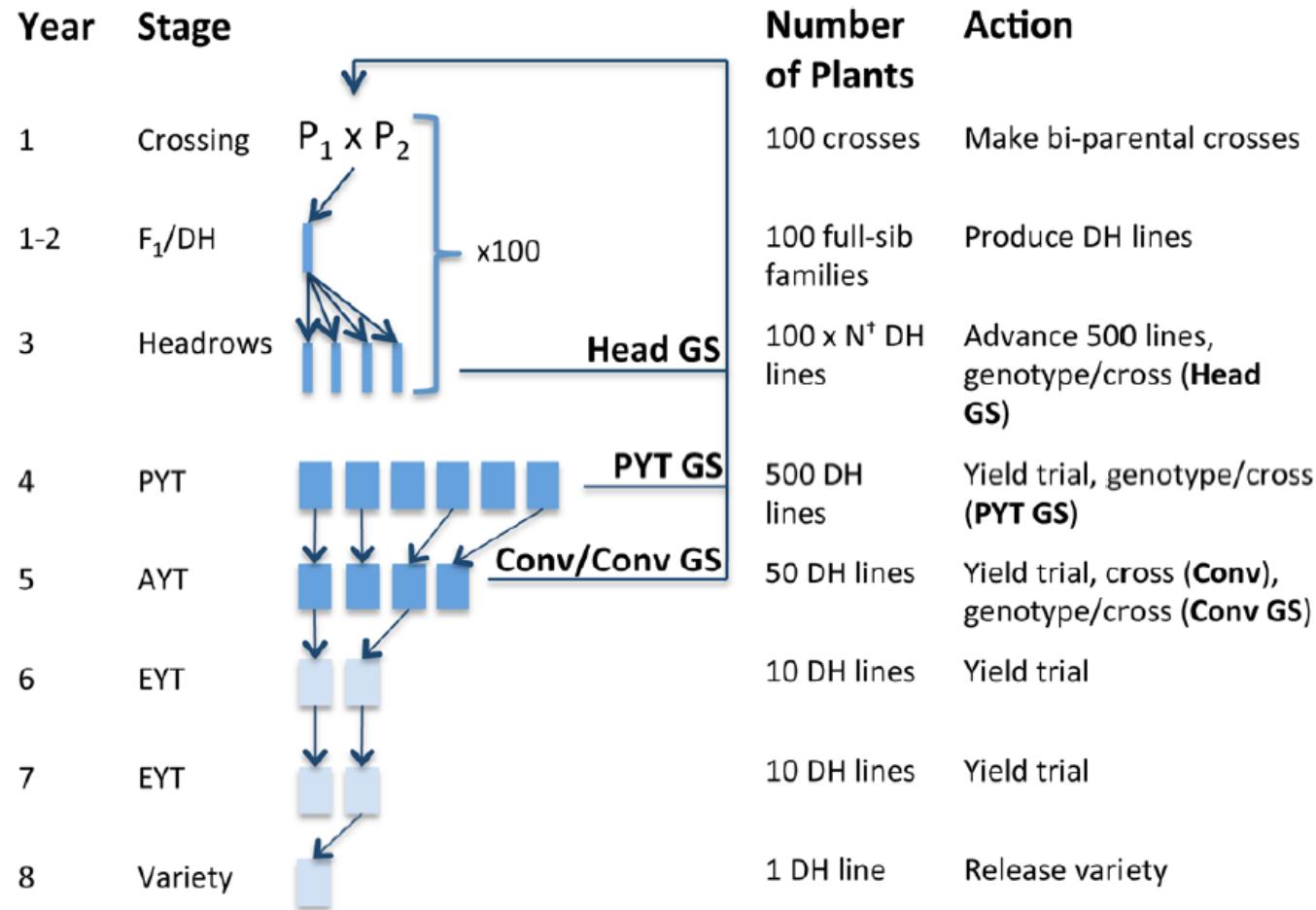
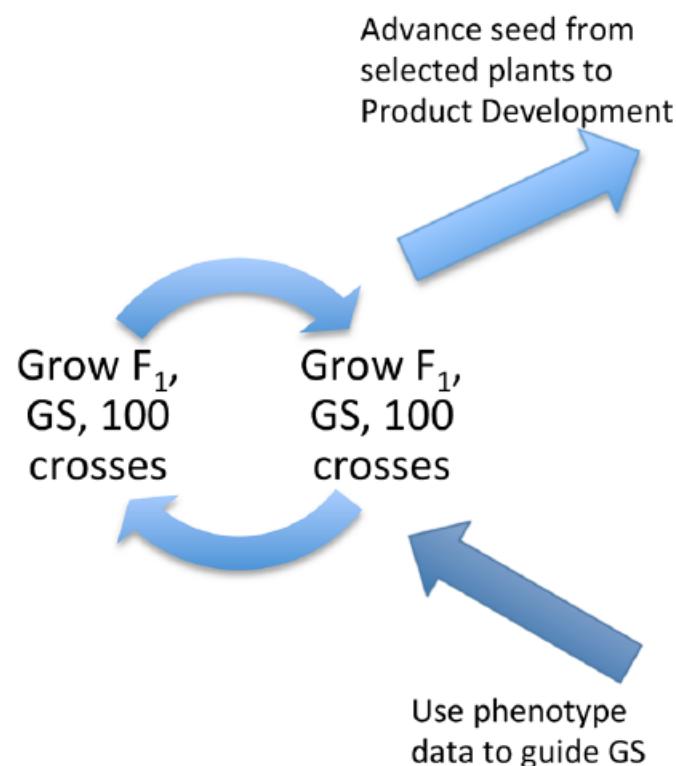


Fig. 1. Overview of breeding schemes for the conventional program (Conv; used in burn-in and as a control) and the programs using standard genomic selection strategies. DH, doubled haploid; PYT, preliminary yield trial; AYT, advanced yield trial; EYT, elite yield trial; Head GS, headrow genomic selection program; PYT GS, preliminary yield trial genomic selection program; Conv GS, conventional program with genomic selection. [†]The number of DH lines per cross (N) differs for each breeding program to maintain equal operating costs. See Table 1.

Gaynor et al (2017) – Figure 2

Population Improvement



Product Development

Year	Stage	Number of Plants	Action
1-2	F_1/DH	200 half-sib families	Produce DH lines
3	Headrow	200 x N^{\dagger} DH lines	Advance 500 lines, genotype (2Part+H)
4	PYT	500 DH lines	Yield trial, genotype (2Part)
5	AYT	50 DH lines	Yield trial
6	EYT	10 DH lines	Yield trial
7	EYT	10 DH lines	Yield trial
8	Variety	1 DH line	Release variety

The product development timeline is shown as a vertical sequence of stages: F_1/DH (Year 1-2), Headrow (Year 3), PYT (Year 4), AYT (Year 5), EYT (Year 6), EYT (Year 7), and Variety (Year 8). Arrows indicate the flow of lines from one stage to the next. In the Headrow stage, 200 lines are advanced to 500 lines. In the PYT stage, 500 lines are reduced to 50 lines. In the AYT stage, 50 lines are reduced to 10 lines. In the EYT stages, 10 lines are further refined until a final variety is released.

Fig. 2. Overview of the two-part program with PYT genomic selection (2Part) and two-part program with headrow genomic selection (2Part+H). DH, doubled haploid; GS, genomic selection; PYT, preliminary yield trial; AYT, advanced yield trial; EYT, elite yield trial. \dagger The number of DH lines per cross (N) differs for each breeding program to maintain equal operating costs. See Table 1.

Gaynor et al (2017) – Figure 3

Simulation Stage	Key Features
Burn-in	Genome Sequence 100,000 generations of evolution Wheat historical effective population size 21 chromosome pairs 1.43 Morgans per chromosome 8×10^8 base pairs per chromosome 2×10^{-9} mutation rate
	Founder Genotypes 50 inbred founders 21,000 SNP markers 21,000 QTN Normally distributed QTN effects
	Recent Breeding 20 years of modern breeding (-19 to 0) Double haploid lines No genomic selection
Evaluation	Future Breeding 20 years of breeding (1 to 20) Testing alternative breeding programs Equal cost programs Ridge regression for genomic selection

Fig. 3. Simulation stages for a single replicate. SNP, single nucleotide polymorphism; QTN, quantitative trait nucleotides.

Gaynor et al (2017) – Figure 4

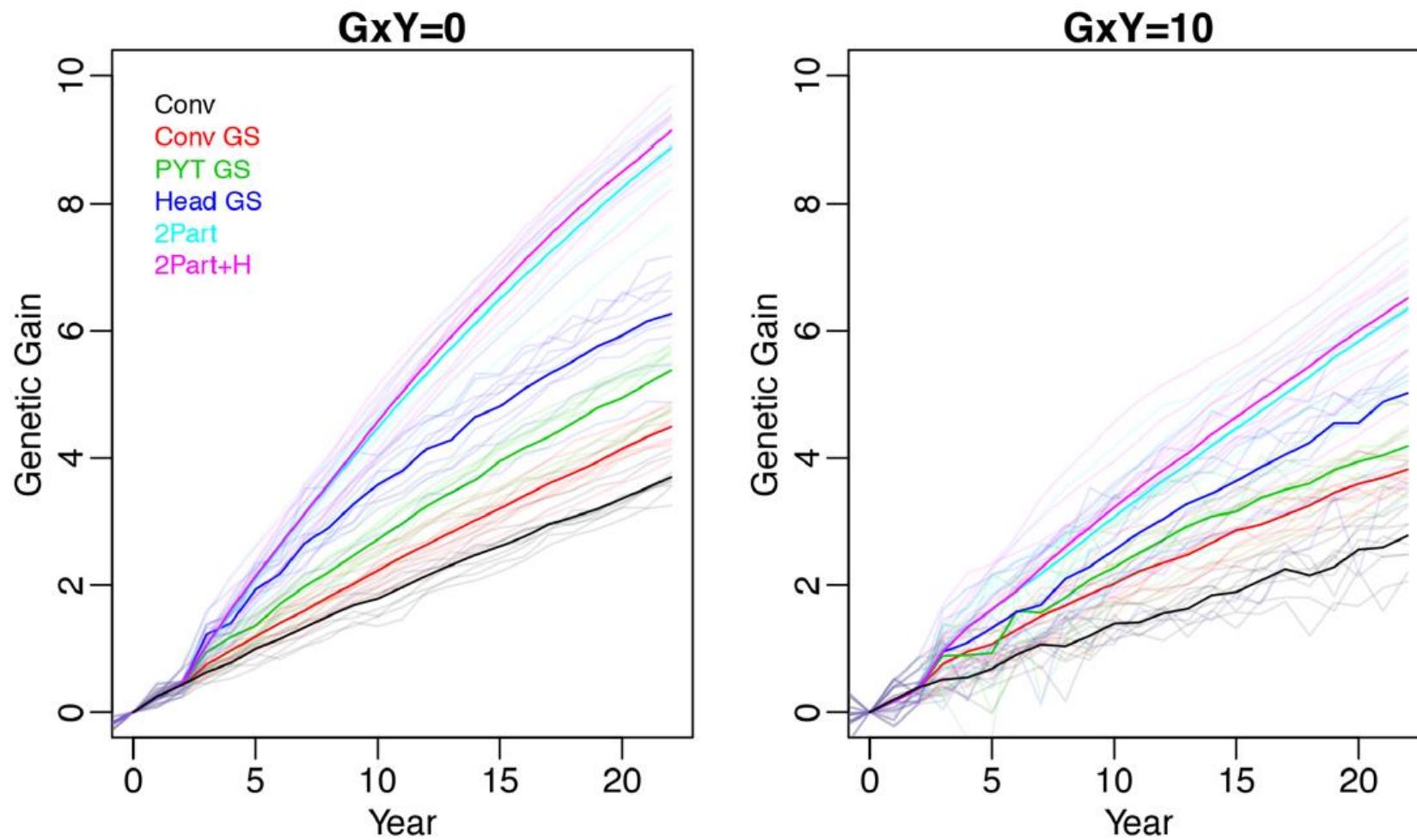


Fig. 4. Genetic gain for all breeding programs when genotype-by-year ($G \times Y$) variance is 0 and 10. Genetic gain is expressed as mean genetic value of headrow entries over time. The mean genetic value for each replicate was centered on 0 in Year 0. Individual replicates are shown with faded lines, and means for all 10 replicates are shown with dark lines. Conv, conventional breeding program; Conv GS, conventional program with genomic selection; PYT GS, preliminary yield trial genomic selection program; Head GS, headrow genomic selection program; 2Part, two-part program with PYT genomic selection; 2Part+H, two-part program with headrow genomic selection.

Gaynor et al (2017) – Figure 5

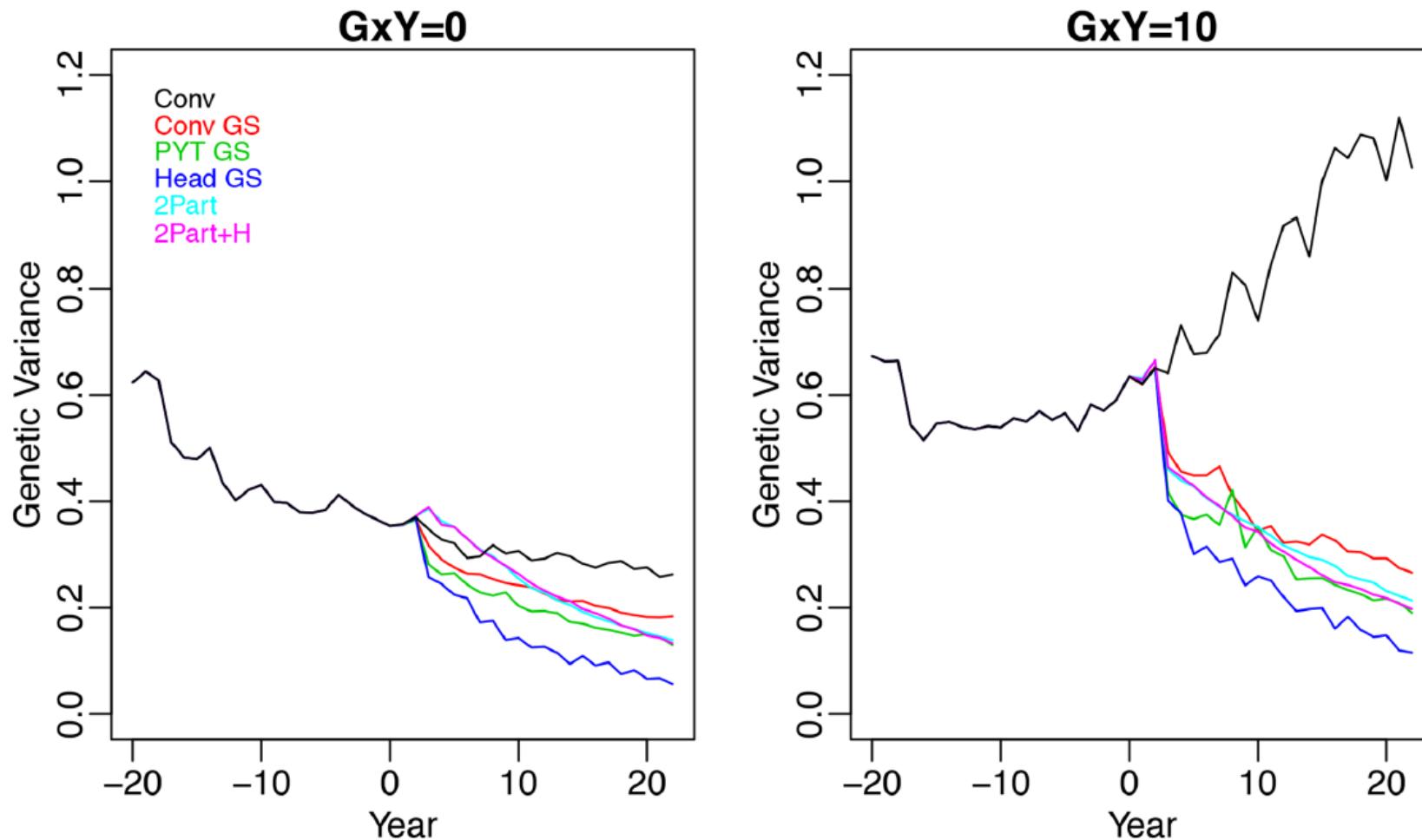


Fig. 5. Genetic variance for all breeding programs when genotype-by-year ($G \times Y$) variance is 0 and 10. Genetic variance is expressed as the genetic variance among headrows in each year of the simulation. Each line represents the overall mean for all 10 replicates. Conv, conventional breeding program; Conv GS, conventional program with genomic selection; PYT GS, preliminary yield trial genomic selection program; Head GS, headrow genomic selection program; 2Part, two-part program with PYT genomic selection; 2Part+H, two-part program with headrow genomic selection.

Gaynor et al (2017) – Figure 6

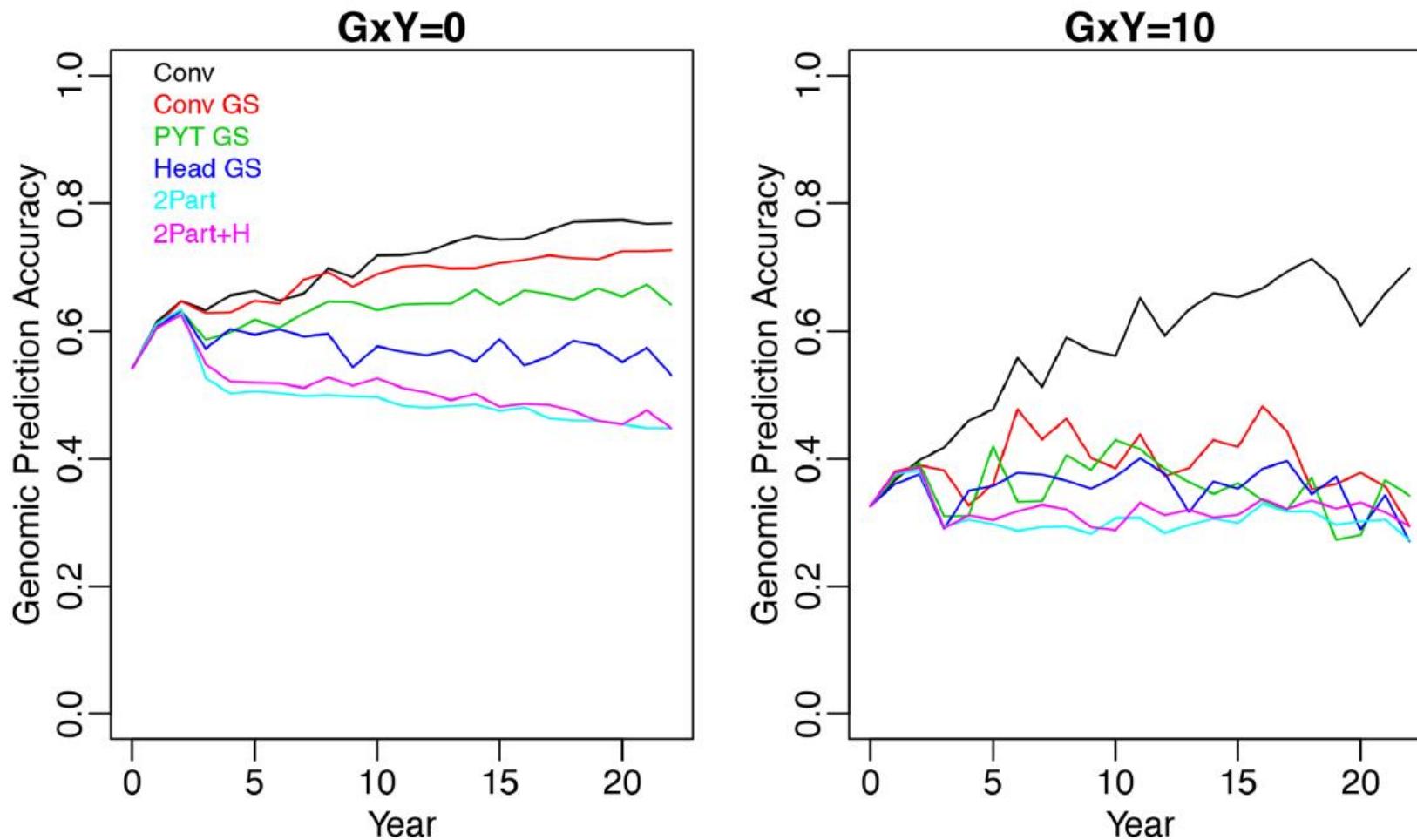


Fig. 6. Genomic prediction accuracy for all breeding programs when genotype-by-year ($G \times Y$) variance is 0 and 10. Genomic prediction accuracy is expressed as the correlation between true and genomic predicted genetic values of headrow entries over time. Conv, conventional breeding program; Conv GS, conventional program with genomic selection; PYT GS, preliminary yield trial genomic selection program; Head GS, headrow genomic selection program; 2Part, two-part program with PYT genomic selection; 2Part+H, two-part program with headrow genomic selection.

Gaynor et al (2017) – Figure 7

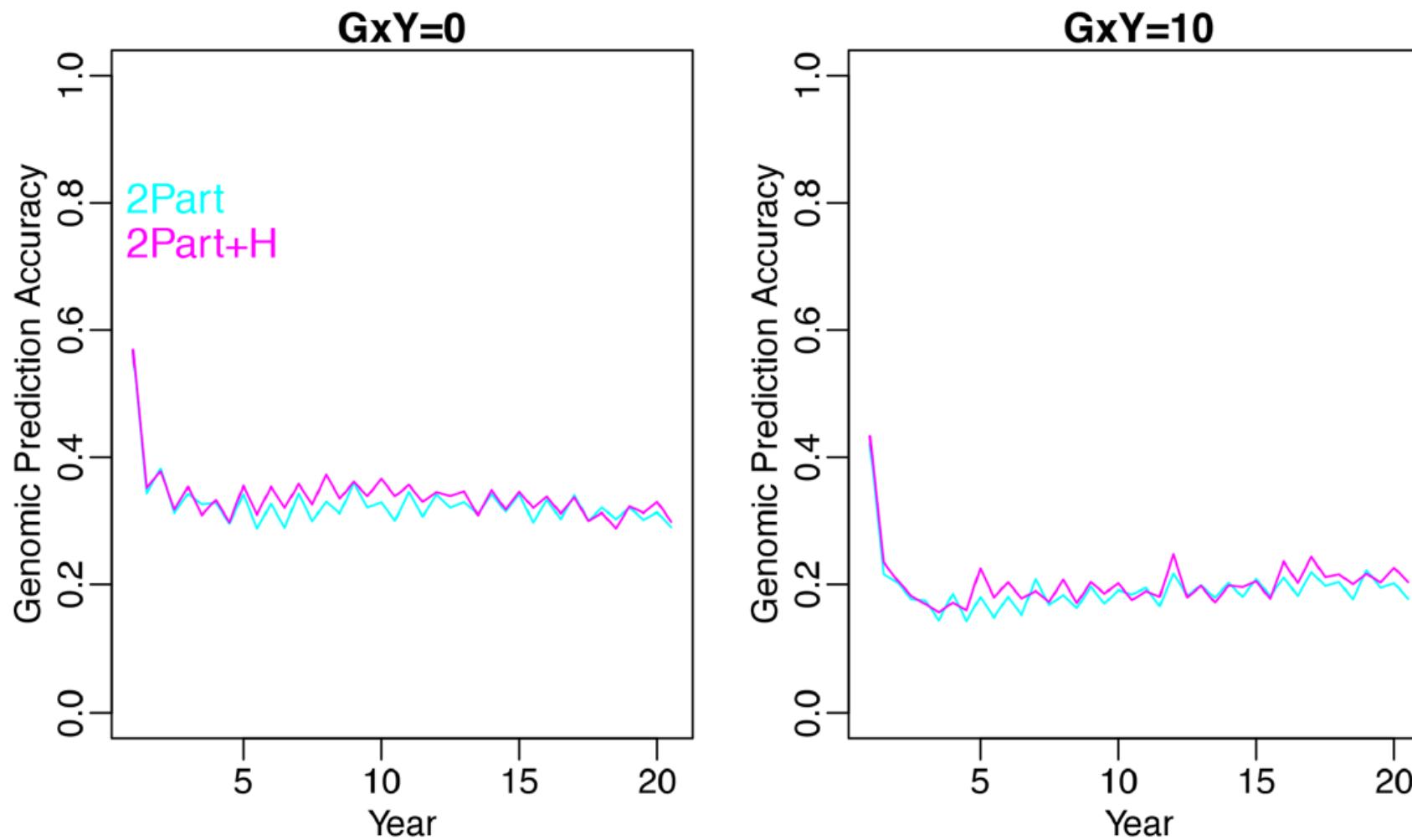


Fig. 7. Genomic prediction accuracy in the population improvement part of the two-part programs when genotype-by-year ($G \times Y$) variance is 0 and 10. Genomic prediction accuracy is expressed as the correlation between true and genomic predicted genetic values of all population improvement plants over time. 2Part, two-part program with PYT genomic selection; 2Part+H, two-part program with headrow genomic selection.