

A photograph of a small, light-colored wooden shed with a dark grey tiled roof. The shed has two doors and is situated in a field of tall, dry grass or crops. The sun is low in the sky, casting long shadows and creating a warm, golden glow over the scene.

Navigating modern plant breeding with quantitative genetics

CJ Yang

Postdoctoral researcher – SRUC

July 24, 2023

Self introduction

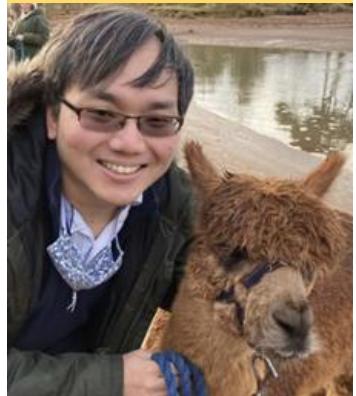
Wisconsin
2012-2018
PhD Genetics



Indiana
2009-2012
BSc Biotech, Math



Edinburgh
2019-now
Postdoc



Freising
2018-2019
Postdoc



Map from R/maps

Malaysia
1990-2009



Talk outline

- Introduction to plant breeding
- Past work
- Current work
- Future work
- Closing
- ???

Introduction – Plant Breeding

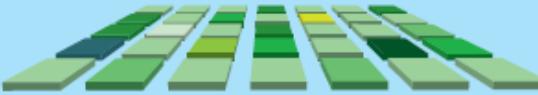
BREEDING 1.0

Incidental selection by farmers



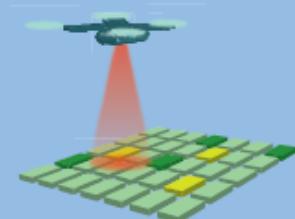
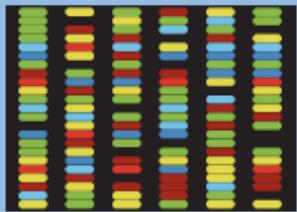
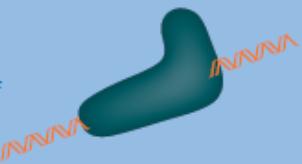
BREEDING 2.0

Statistical and experimental design to improve selection effort



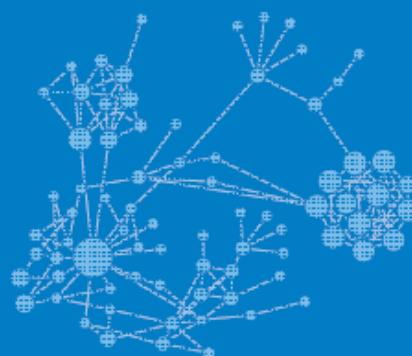
BREEDING 3.0

Integration of genetic and genomic data; current state of the art



BREEDING 4.0

Ability to combine any known alleles into optimal combinations; will be reached soon for some crops



Wallace et al (2018)

Introduction – Plant Breeding

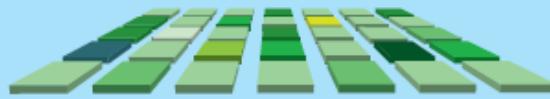
BREEDING 1.0

Incidental selection by farmers



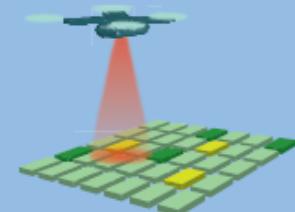
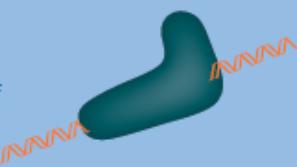
BREEDING 2.0

Statistical and experimental design to improve selection effort



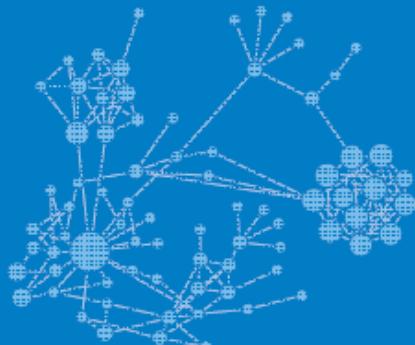
BREEDING 3.0

Integration of genetic and genomic data; current state of the art



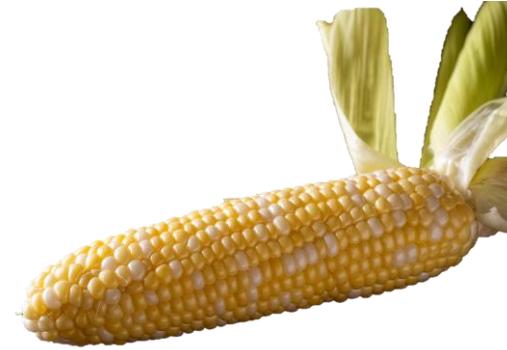
BREEDING 4.0

Ability to combine any known alleles into optimal combinations; will be reached soon for some crops



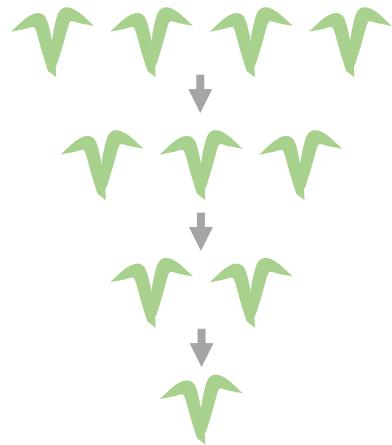
Wallace et al (2018)

Different crops are at different stages



Introduction – Breeding Methods

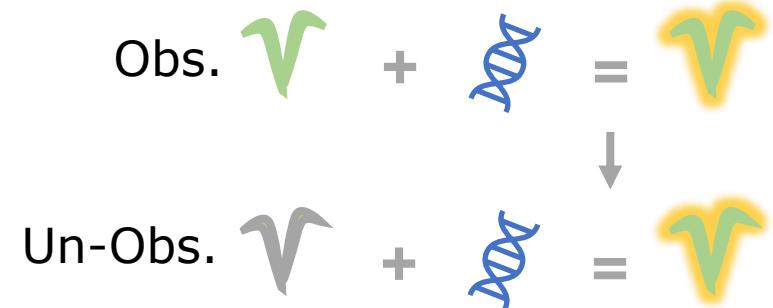
Pedigree breeding



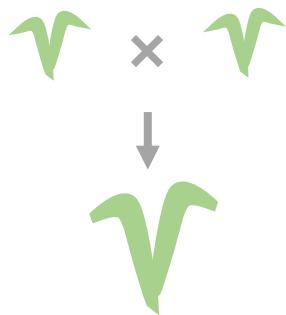
Mutation breeding



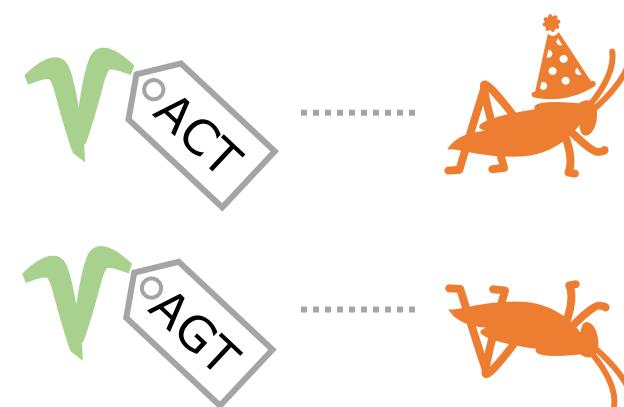
Genomic selection (GS)



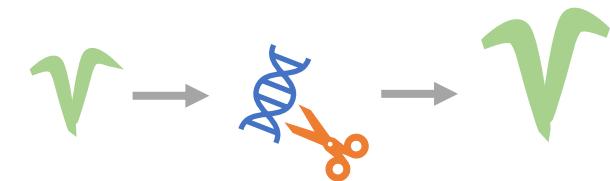
Hybrid breeding



Marker assisted selection (MAS)

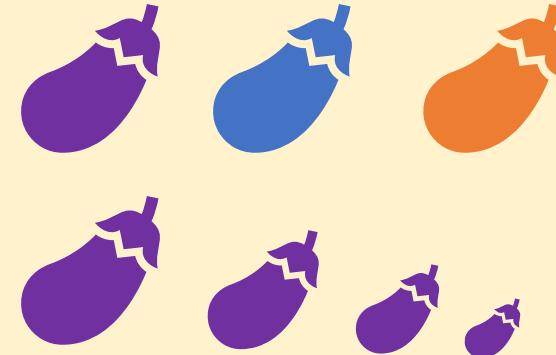


Gene editing (GE)



Introduction – Breeding Methods

Trait genetic architecture



Time and cost

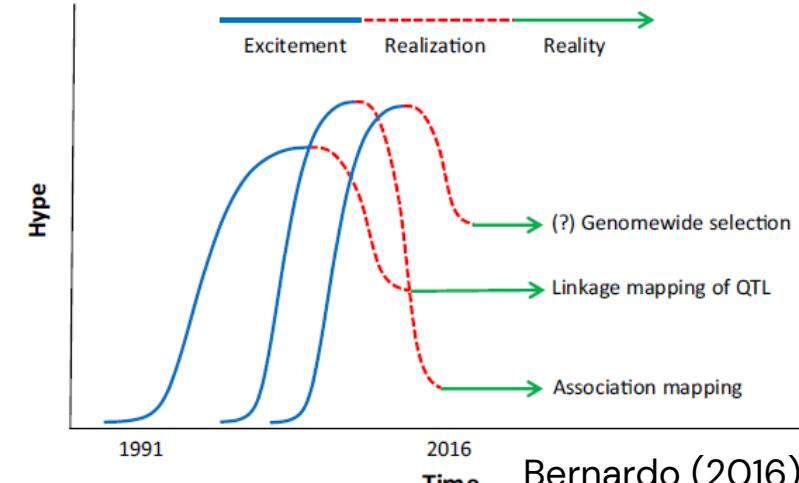


Genetic diversity

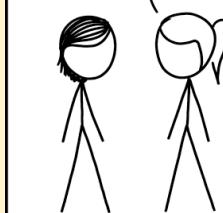


https://waldo.fandom.com/wiki/File:Received_2652948628312295.jpeg

Hype



ALL READY FOR THE
HEARING? LET'S GO OVER
THINGS ONE MORE TIME.
AHEM
WHAT IS THE MAIN REASON
TO FUND THIS MISSION?



IT WILL SIGNIFICANTLY ADVANCE OUR LONG-TERM GOAL OF BETTER UNDERSTANDING
THE FORMATION AND EVOLUTION OF THE
SOLAR SYSTEM, WHILE FULFILLING OUR
MANDATE TO DEVELOP A NEW GENERATION
OF INTERPLANETARY SPACECRAFT.



AND BECAUSE
IT'S SPACE!
SPAAAAACE.
PEW PEW PEW!
SPACE!



<https://xkcd.com/2124>

Introduction – Genetic Gain

Breeder's equation
(Lush 1937)

$$R = h^2 S$$

$$R = \frac{\sigma_g^2}{\sigma_p^2} \sigma_p i$$

$$R = \frac{\sigma_g}{\sigma_p} \sigma_g i$$

$$R = h\sigma_g i$$

Rate of genetic gain

$$\Delta R = \frac{h\sigma_g i}{t}$$

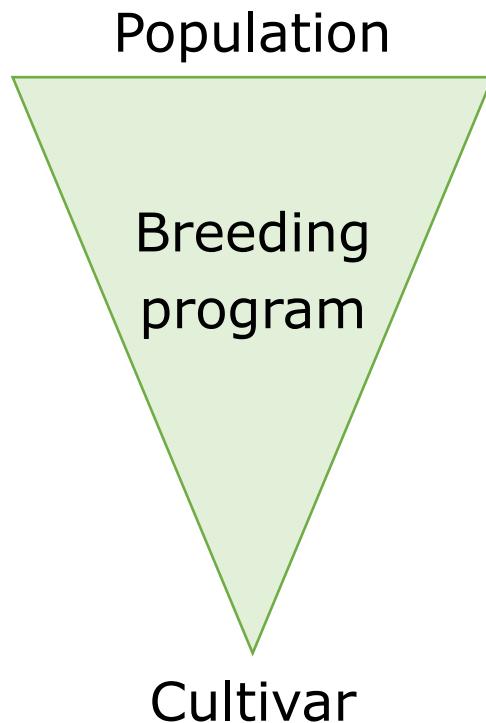
$$\Delta R = \frac{\text{sel. accuracy} \times \text{genetic variation} \times \text{sel. intensity}}{\text{time}}$$



Framework for quantitative genetics

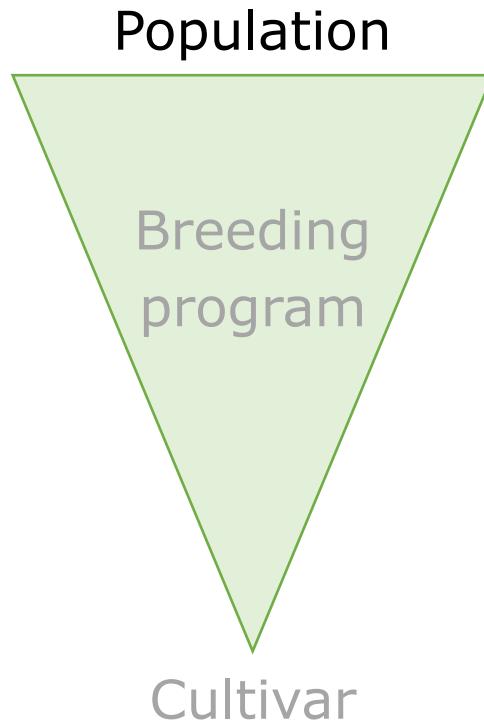
Introduction – Improving ΔR

$$\Delta R = \frac{\text{sel. accuracy} \times \text{genetic variation} \times \text{sel. intensity}}{\text{time}}$$



Introduction – Improving ΔR

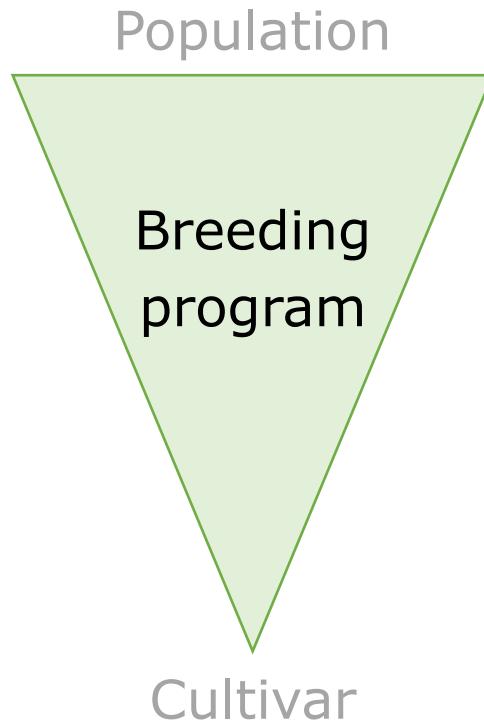
$$\Delta R = \frac{\text{sel. accuracy} \times \text{genetic variation} \times \text{sel. intensity}}{\text{time}}$$



Begin with a high genetic diversity – MPP, mutation, pre-breeding.

Introduction – Improving ΔR

$$\Delta R = \frac{\text{sel. accuracy} \times \text{genetic variation} \times \text{sel. intensity}}{\text{time}}$$

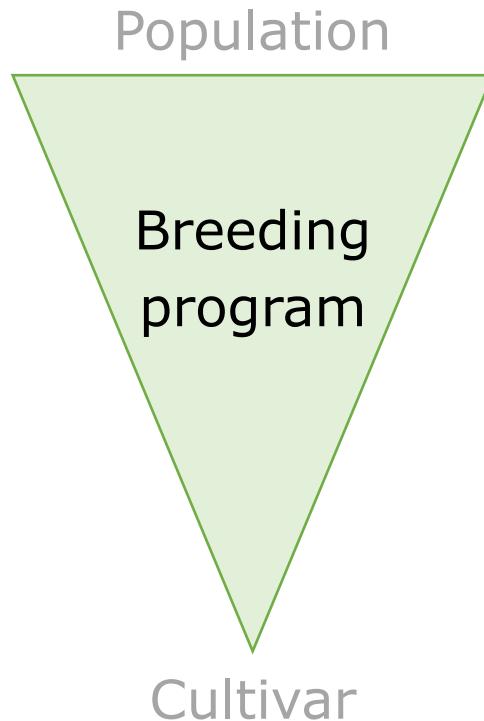


Begin with a high genetic diversity – MPP, mutation, pre-breeding.

Improve selection accuracy – MAS, GS, GE, phenomics, trials.

Introduction – Improving ΔR

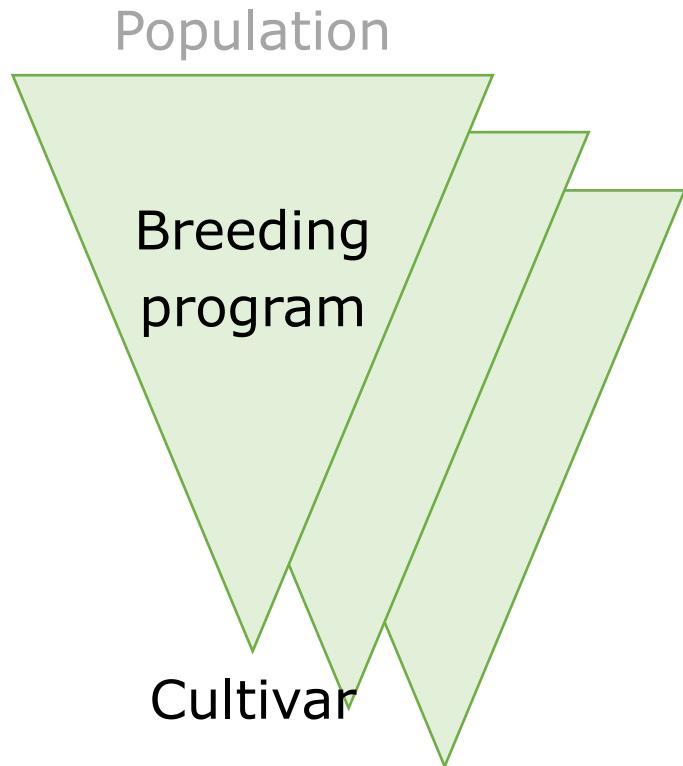
$$\Delta R = \frac{\text{sel. accuracy} \times \text{genetic variation} \times \text{sel. intensity}}{\text{time}}$$



- Population
 - Begin with a high genetic diversity – MPP, mutation, pre-breeding.
- Breeding program
 - Improve selection accuracy – MAS, GS, GE, phenomics, trials.
- Cultivar
 - Increase selection intensity – larger trials, cheaper evaluation.

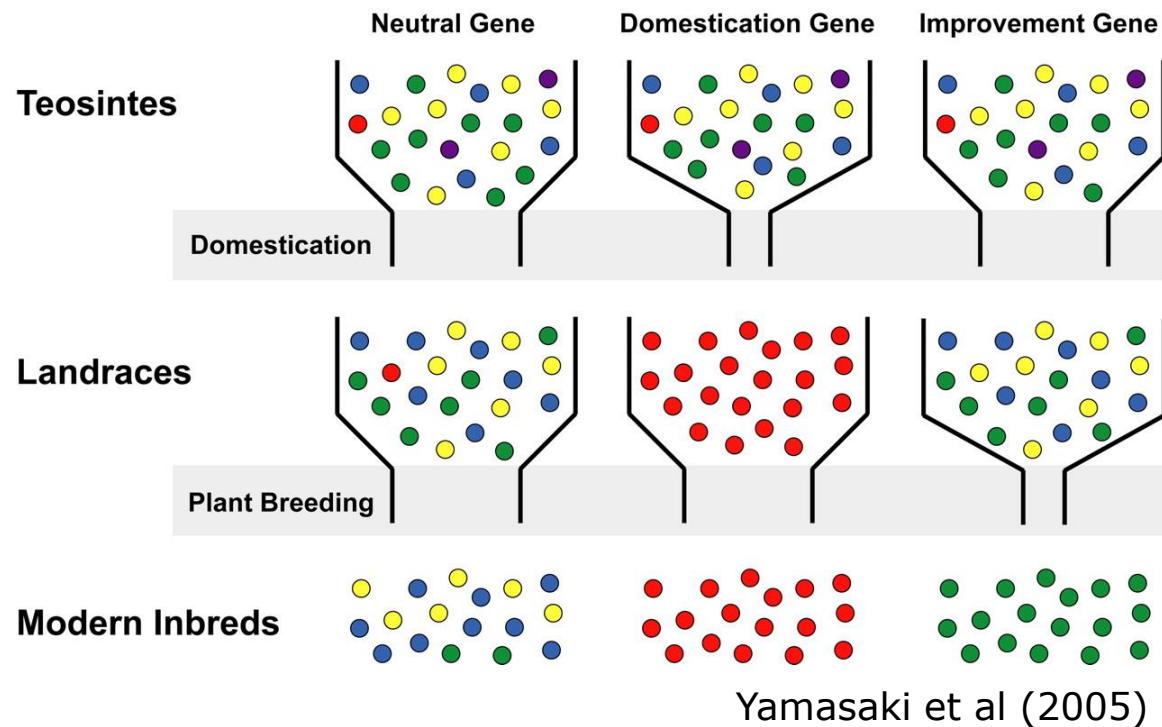
Introduction – Improving ΔR

$$\Delta R = \frac{\text{sel. accuracy} \times \text{genetic variation} \times \text{sel. intensity}}{\text{time}}$$



- Population
 - Begin with a high genetic diversity – MPP, mutation, pre-breeding.
- Breeding program
 - Improve selection accuracy – MAS, GS, GE, phenomics, trials.
 - Increase selection intensity – larger trials, cheaper evaluation.
- Cultivar
 - Reduce breeding cycle time – rapid cycling, SB/RGA.

Shuffling genetic diversity

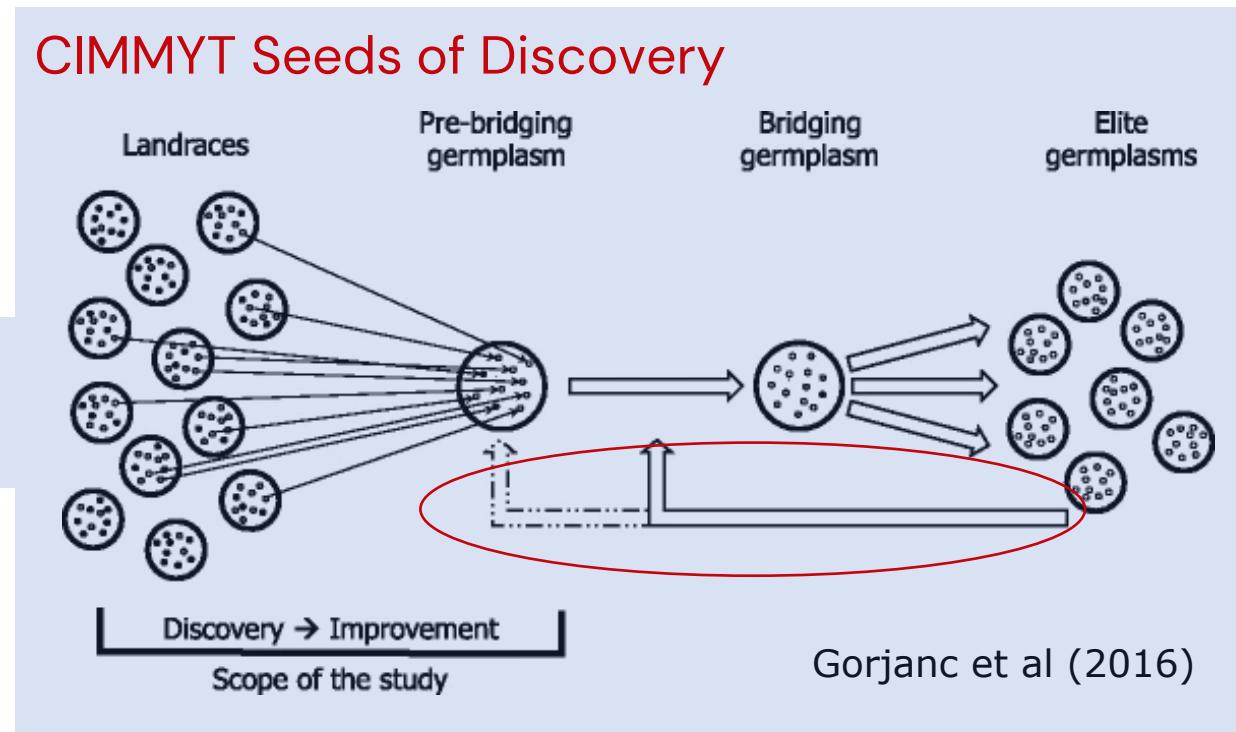
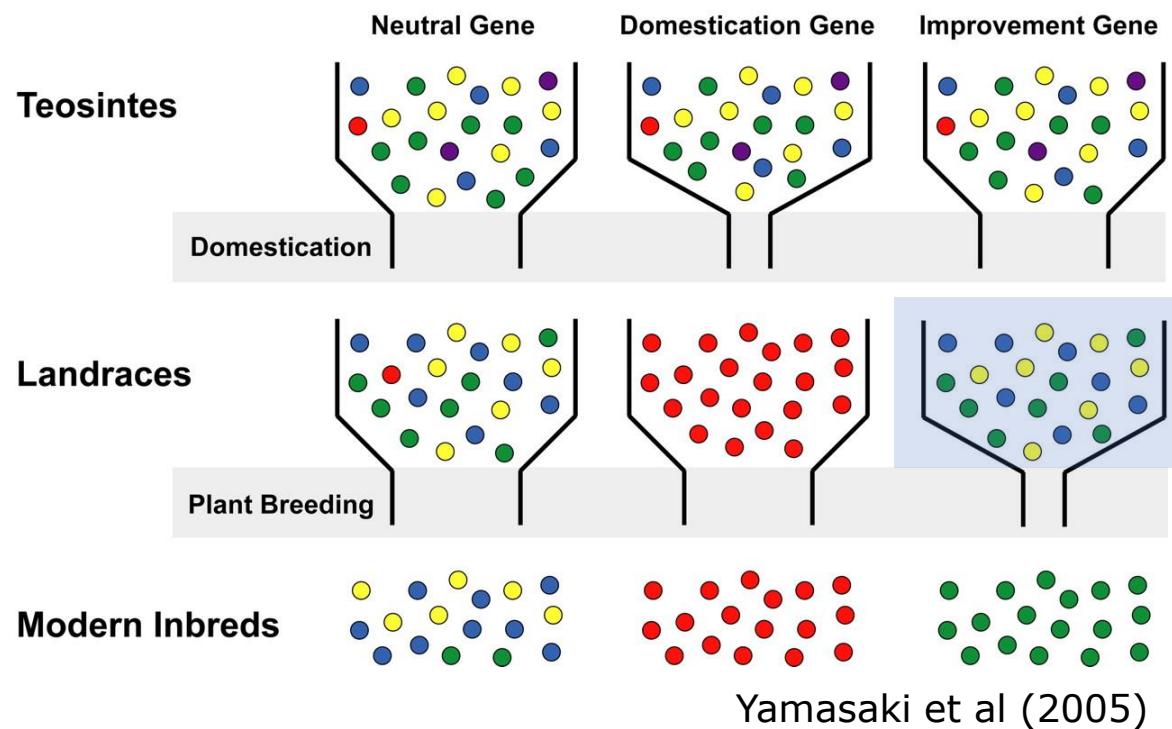


Selection reduces genetic diversity.

Rely on exotic sources (wild/landrace) for novel alleles, e.g. stress tolerance/resistance.

Linkage drag is a challenge.

Shuffling genetic diversity

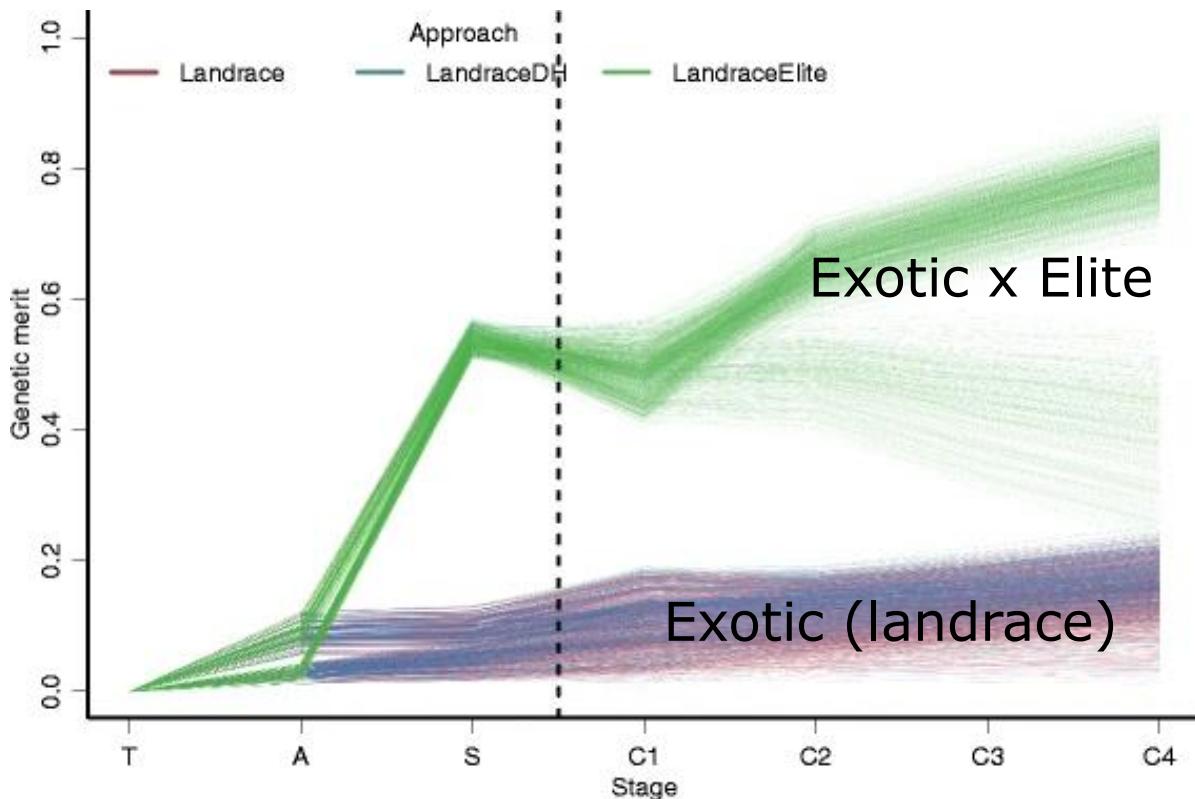


Selection reduces genetic diversity.

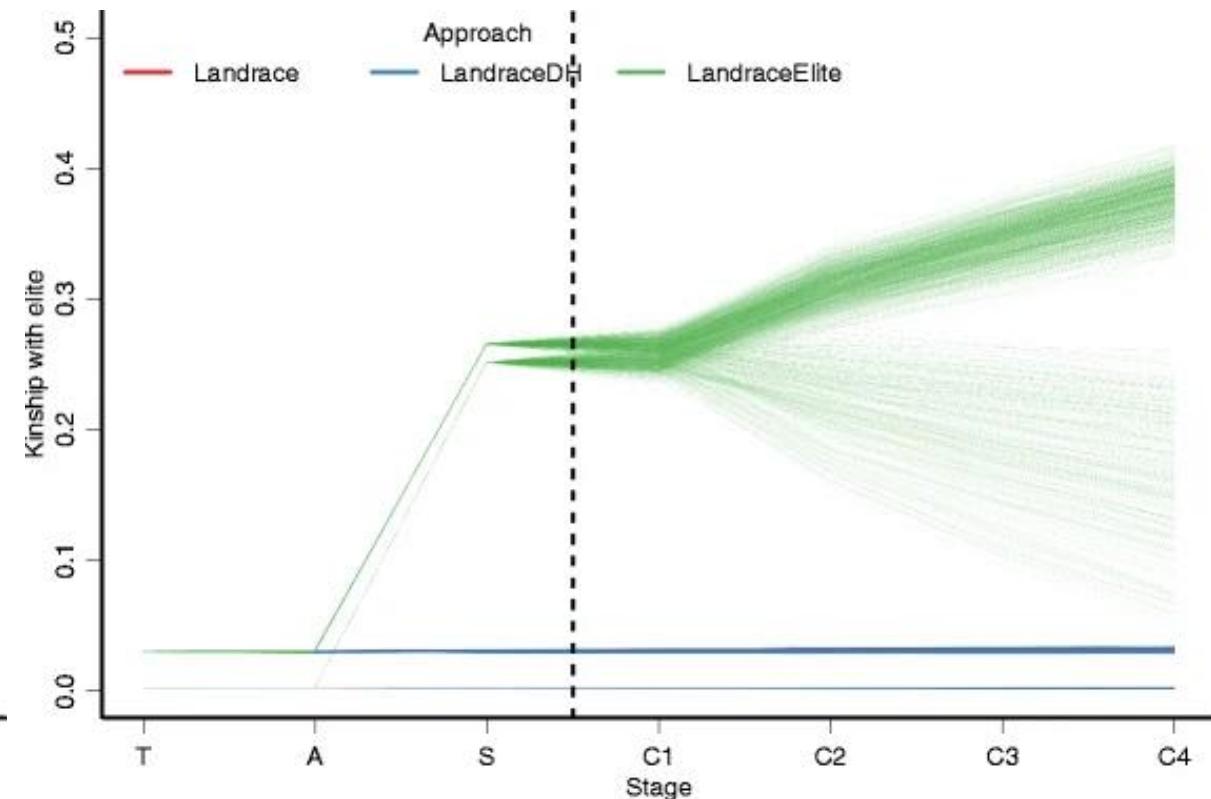
Rely on exotic sources (wild/landrace) for novel alleles, e.g. stress tolerance/resistance.

Linkage drag is a challenge.

Selection bias



Selection within exotic is slow.

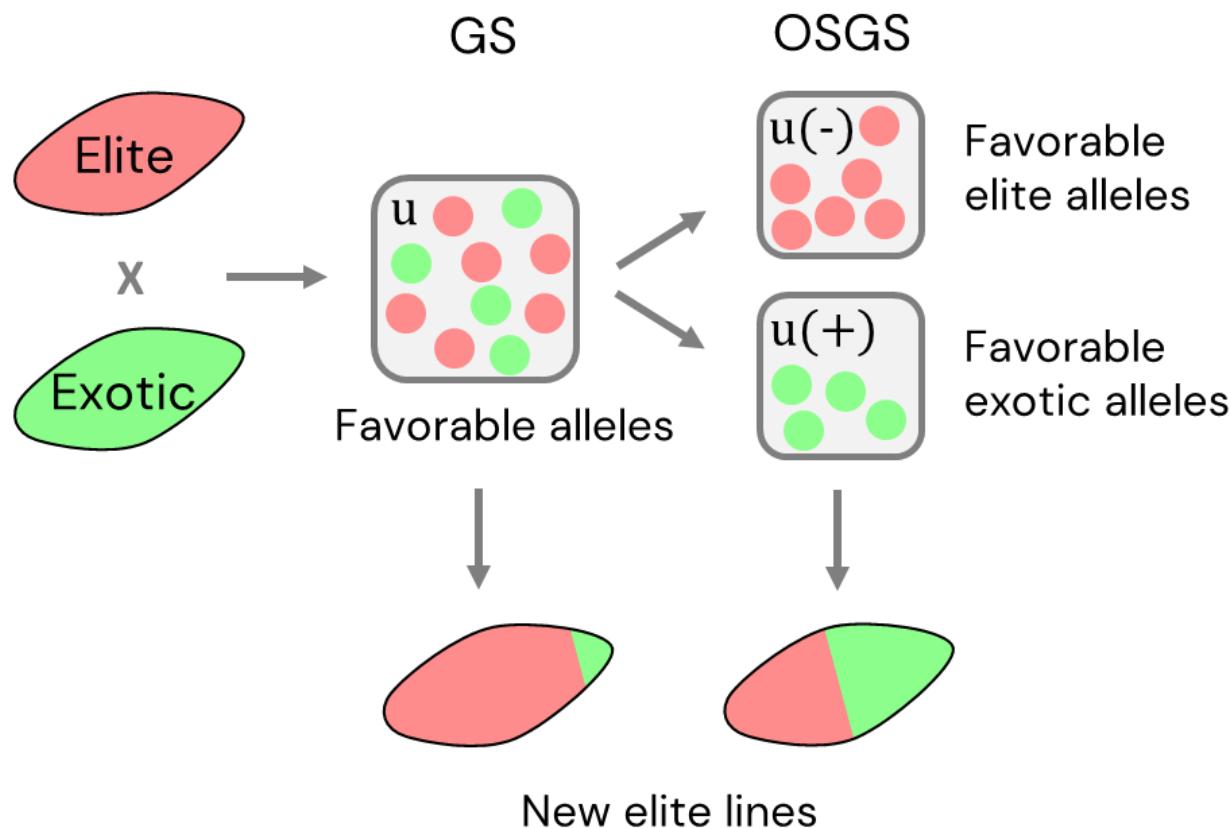


Selection within exotic x elite reconstitutes the elite genome.

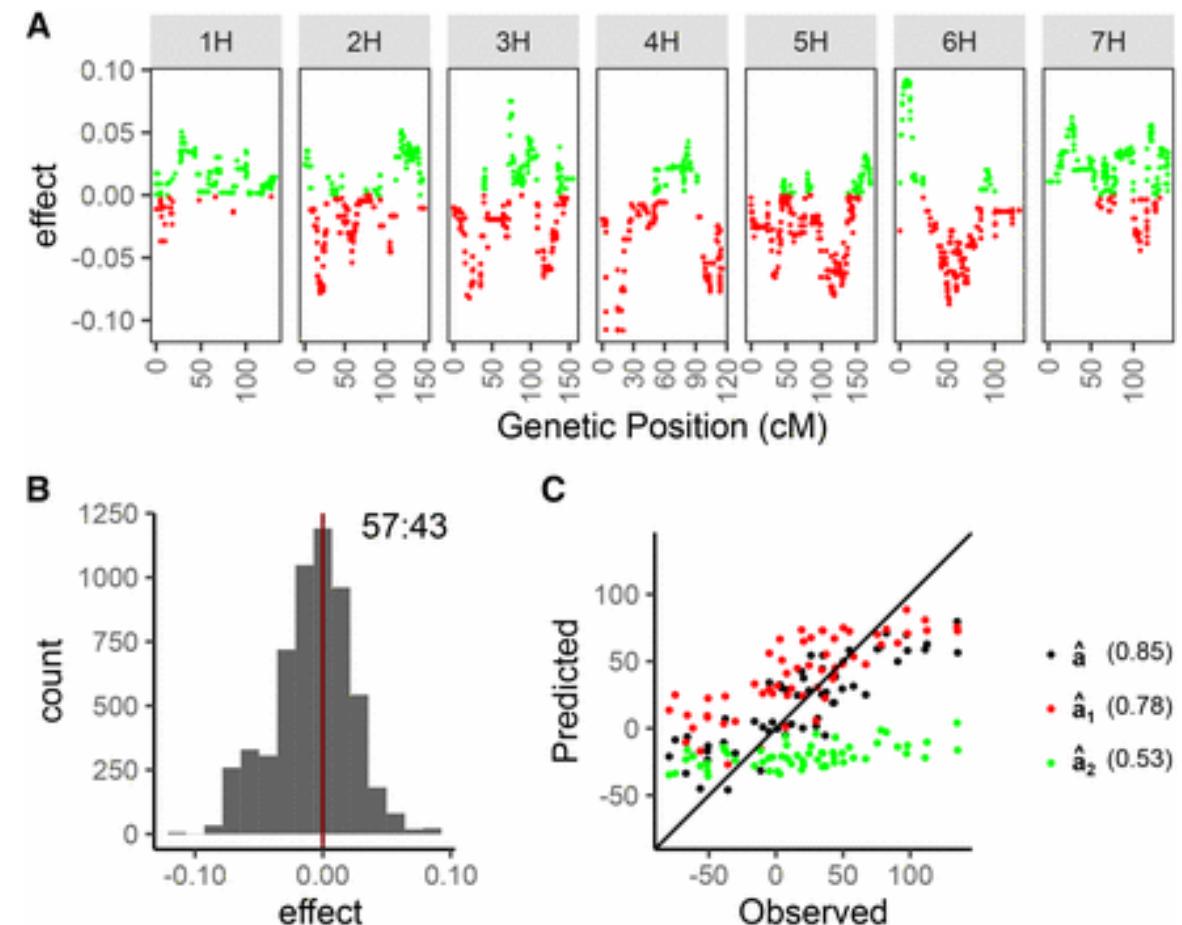
Origin Specific Genomic Selection (OSGS)

How do we improve selection in exotic x elite?

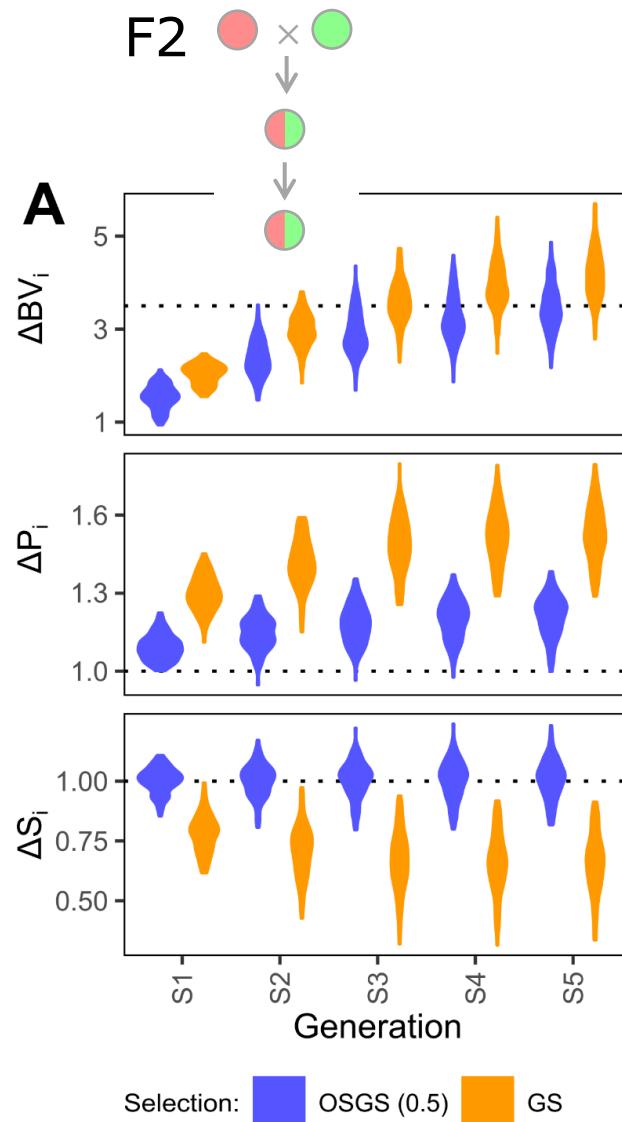
Can we target favorable exotic alleles?



Yield in a barley NAM family (BC1)

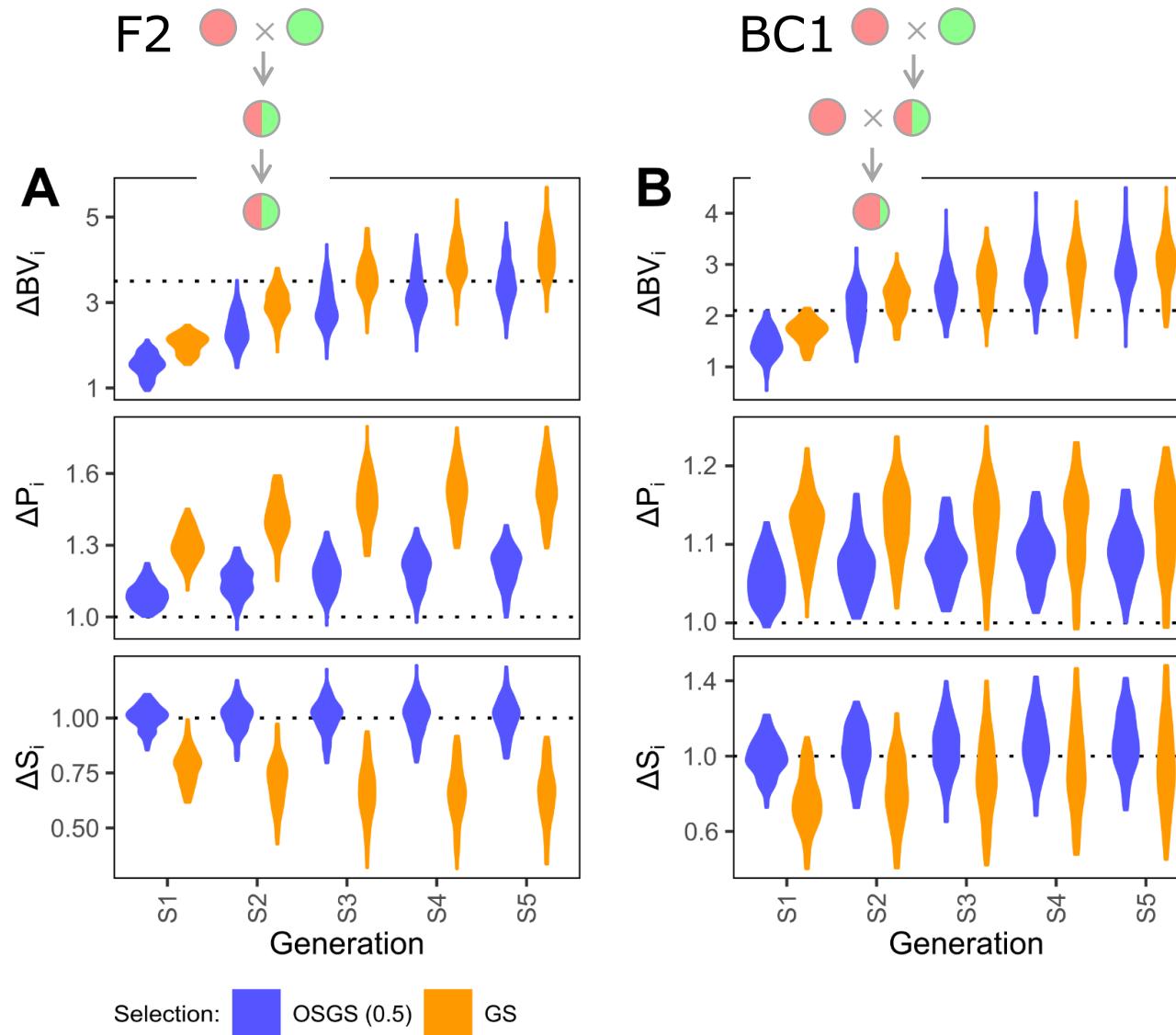


Origin Specific Genomic Selection (OSGS)



- Simulation with 60:40 split in favorable alleles between the elite and exotic parents.
- OSGS prevents selection on the favorable alleles from the elite parent.

Origin Specific Genomic Selection (OSGS)



- Simulation with 60:40 split in favorable alleles between the elite and exotic parents.
- OSGS prevents selection on the favorable alleles from the elite parent.
- The advantage of OSGS diminishes in BC population.

Well, CIMMYT did it anyway...



Direct introgression of untapped diversity into elite wheat lines



<https://knowyourmeme.com/memes/dissappointed-muhammad-sarim-akhtar>

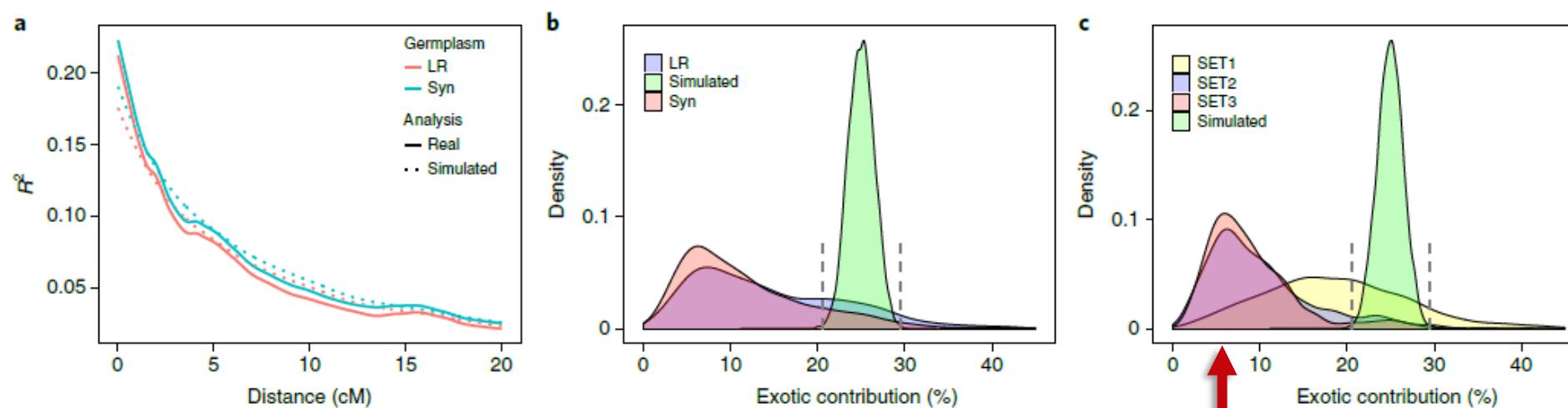
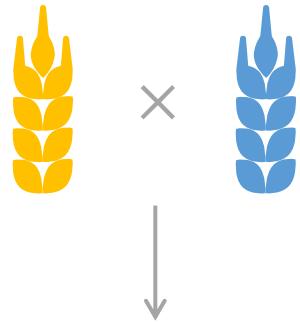


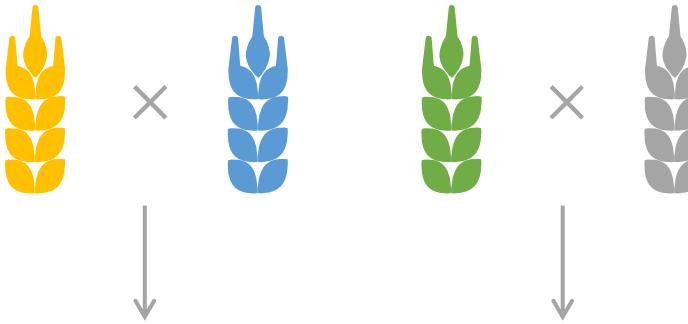
Fig. 1 | Simulation of the pre-breeding germplasm. **a**, The LD decay of empirical and simulated PBLs with Syn and LR backgrounds. **b**, Density plot of the percentage of exotic contribution (LR in blue and Syn in red) to PBLs compared with the simulated expected contribution (green) under neutrality with no selection. **c**, Density plot of the percentage of exotic contribution (SET1 in yellow, SET2 in blue and SET3 in red) to PBLs compared with the simulated expected contribution (green) under neutrality with no selection. Dashed lines in **b** and **c** represent the 95% confidence interval for the expected exotic contribution.

Multiparental population design

Bi-Parental Population (BPP)



Multi-Parental Population (MPP)

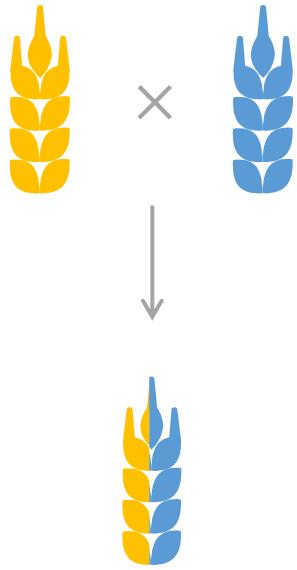


MPP has higher genetic diversity
than BPP.

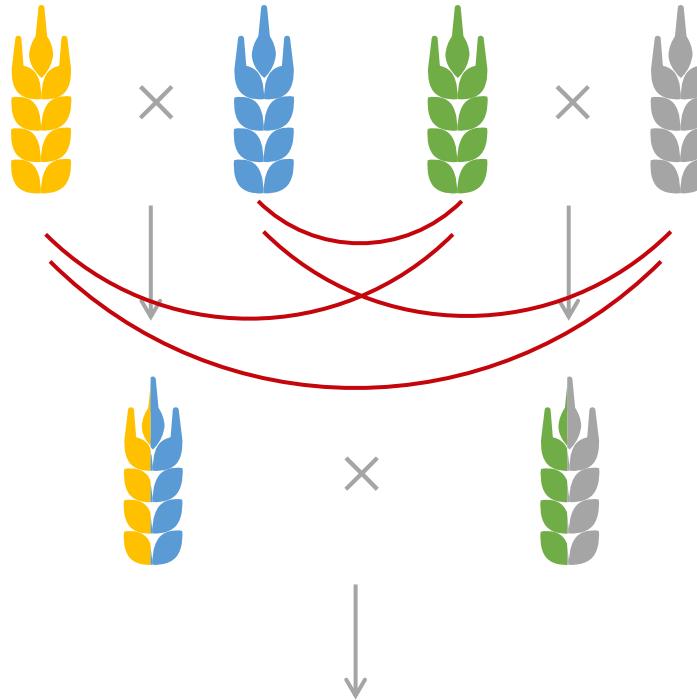


Multiparental population design

Bi-Parental Population (BPP)



Multi-Parental Population (MPP)



MPP has higher genetic diversity than BPP.

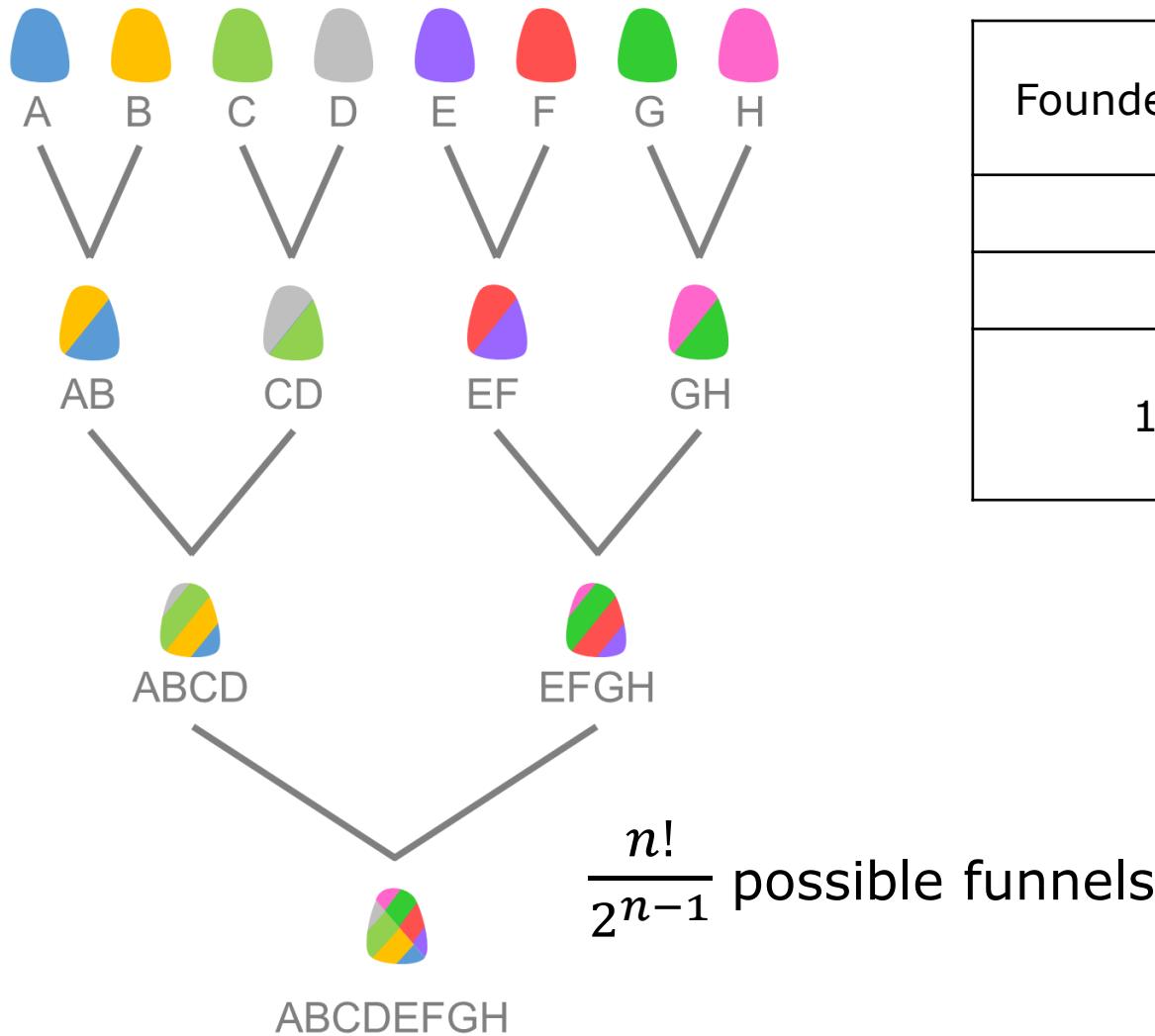


MAGIC: Multiparental Advanced Generation Inter Cross

Multi-generation diallel

Funnel

MAGIC crossing scheme



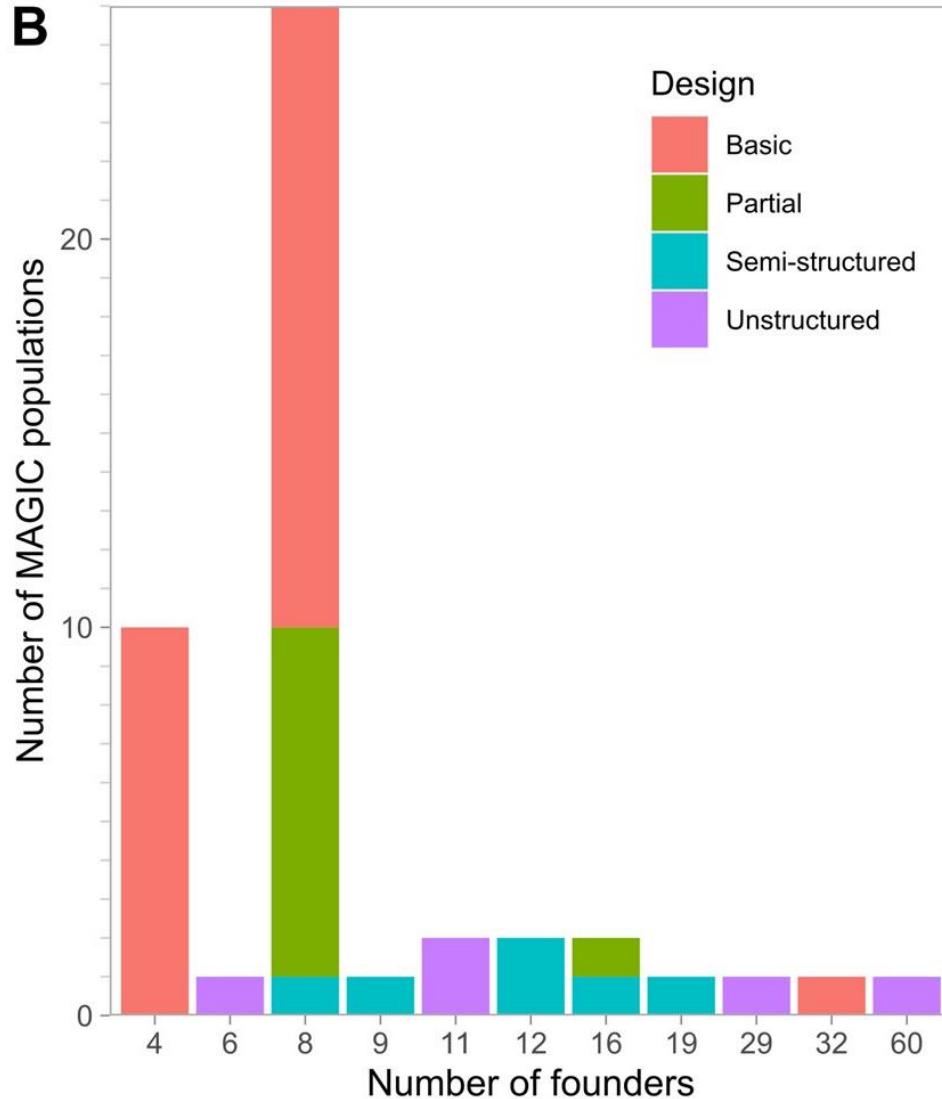
Founder	Basic		Full	
	Funnel	Cross	Funnel	Cross
4	1	2+1	3	6 + 3
8	1	4+2+1	315	28 + 210 + 315
16	1	8+4+2+1	638,512,875	120 + 5,460 + 4,054,050 + 638,512,875

Too few & biased

Too many

MAGIC crossing scheme

B



Founder	Basic		Full	
	Funnel	Cross	Funnel	Cross
4	1	2+1	3	6 + 3
8	1	4+2+1	315	28 + 210 + 315
16	1	8+4+2+1	638,512,875	120 + 5,460 + 4,054,050 + 638,512,875

Too few & biased

Too many

Many MAGIC populations are poorly designed.

R/magicdesign

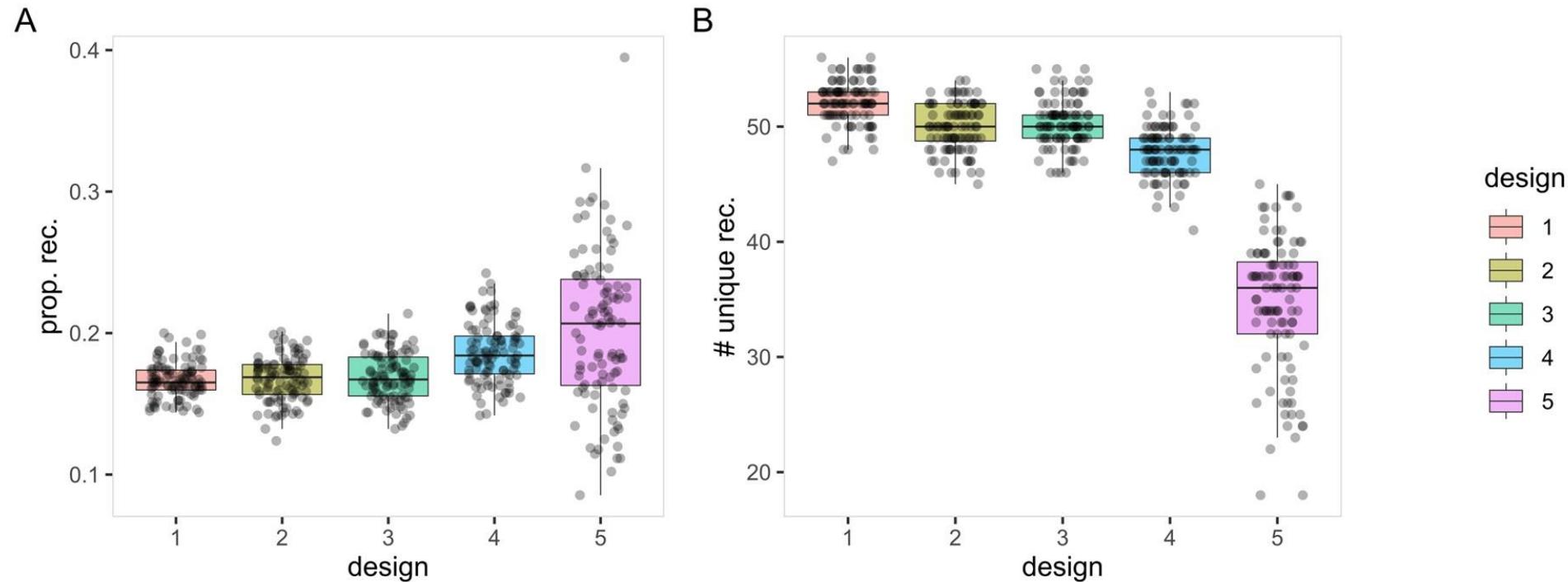
R/magicdesign – create and test via simulation.

Shiny/magicdesignee – web/local GUI version.



More information at
<https://cjyang-work.github.io/>

Example: compare the recombination profiles in 5 different designs.



Alternative GWAS

Usually, we use this model for GWAS.

$$Trait = SNP + Fixed + Random + Residual$$

Fixed effects can be year, location, principal components (PC).

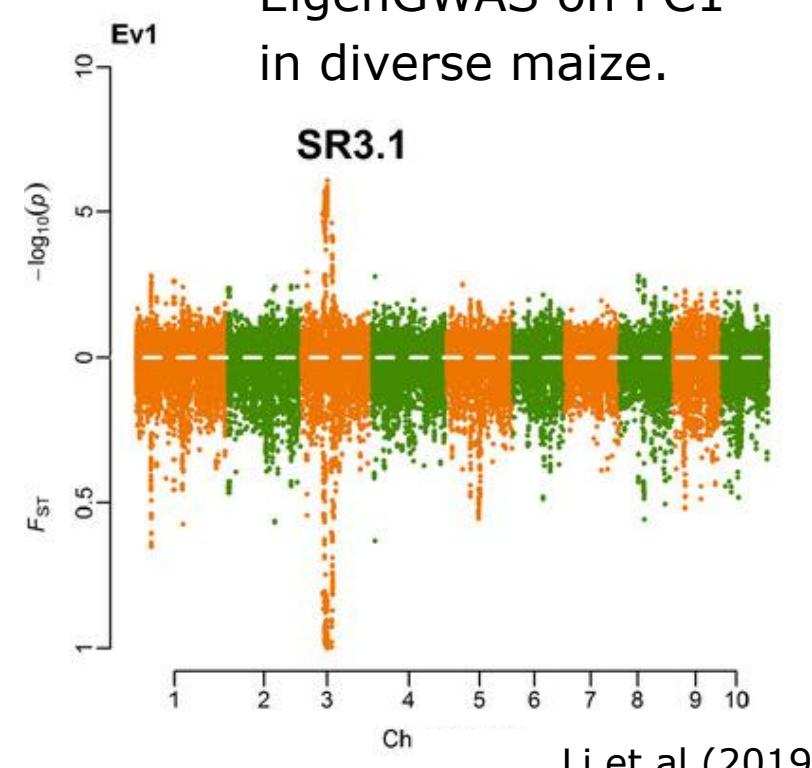
Random effects can be genetic (polygenic) background.

In the absence of trait data, we can do the following.

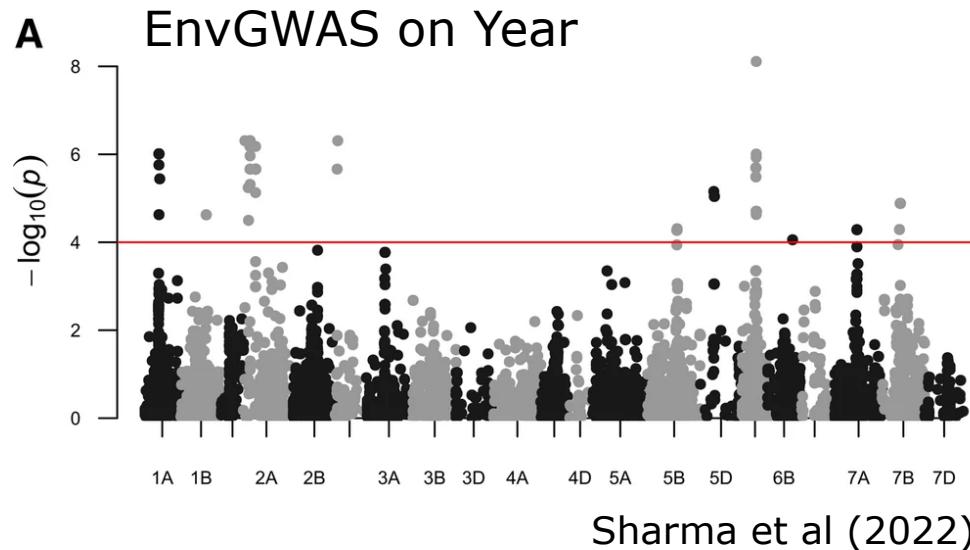
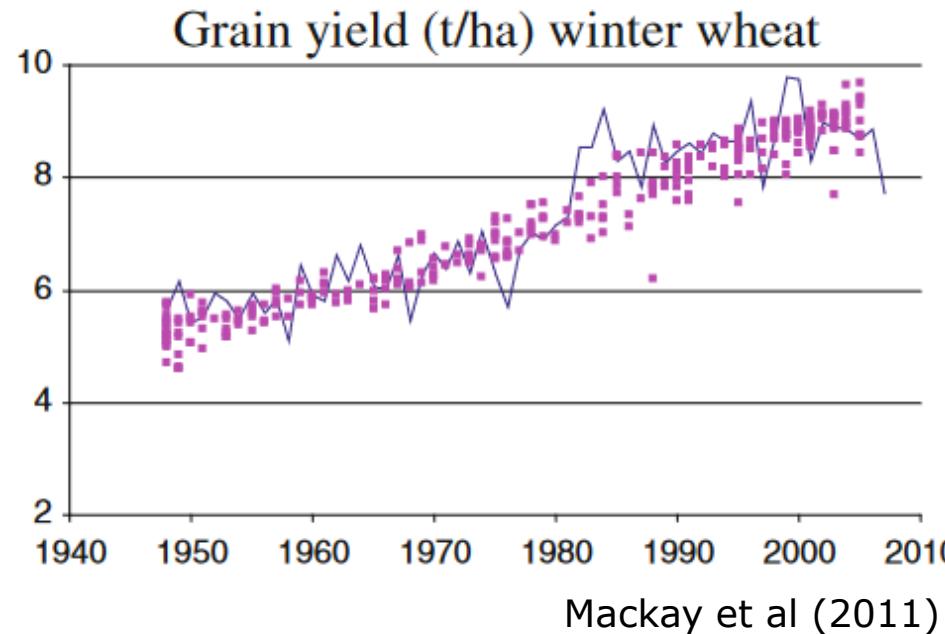
$$Year = SNP + Random + Residual$$

$$PC = SNP + Random + Residual$$

The idea here is that we are mapping QTL for traits that are correlated to Year/PC – **env/eigenGWAS**



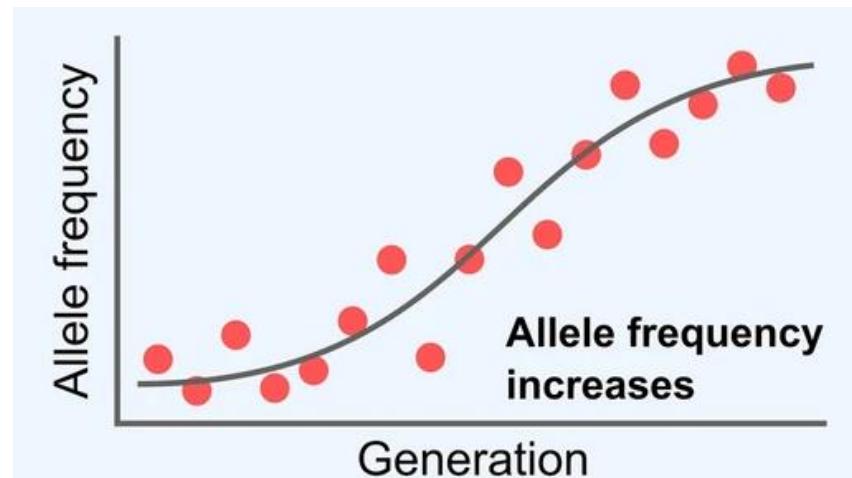
Selection Mapping



Instead of $Year \sim SNP$,
how about Allele frequency $\sim Year$?

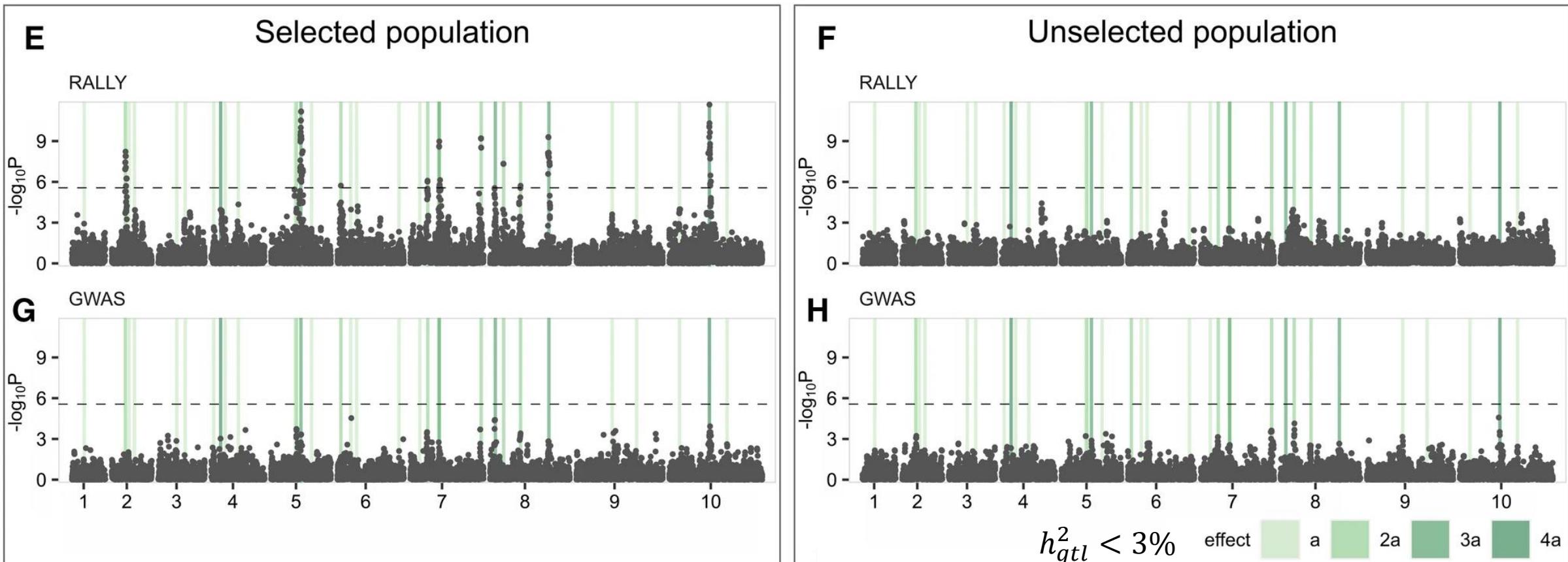
For each marker, this can be done
via a logistic regression model.

$$p_i = \frac{1}{1 + e^{-(\mu_i + \beta_i X)}}$$



RALLY: Regression of ALLeles on Years

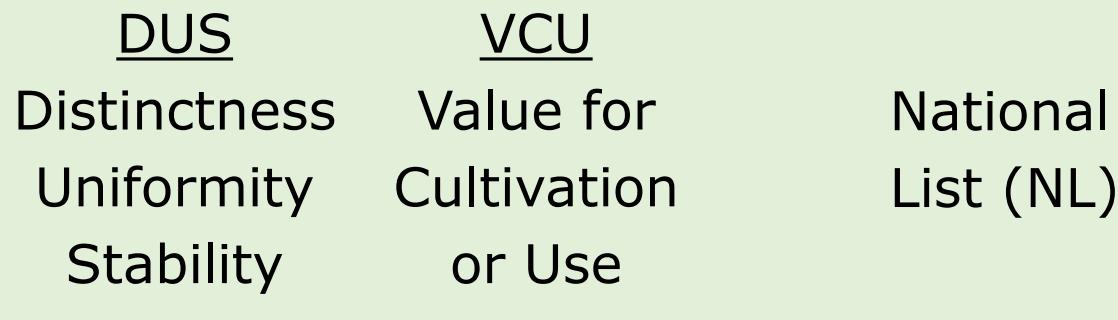
RALLY is good for mapping loci under selection.



False positive control via Parametric Control (Delta + Genomic Control)

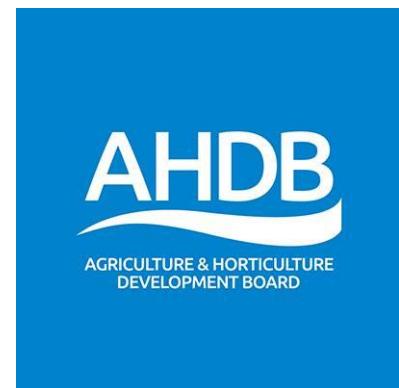
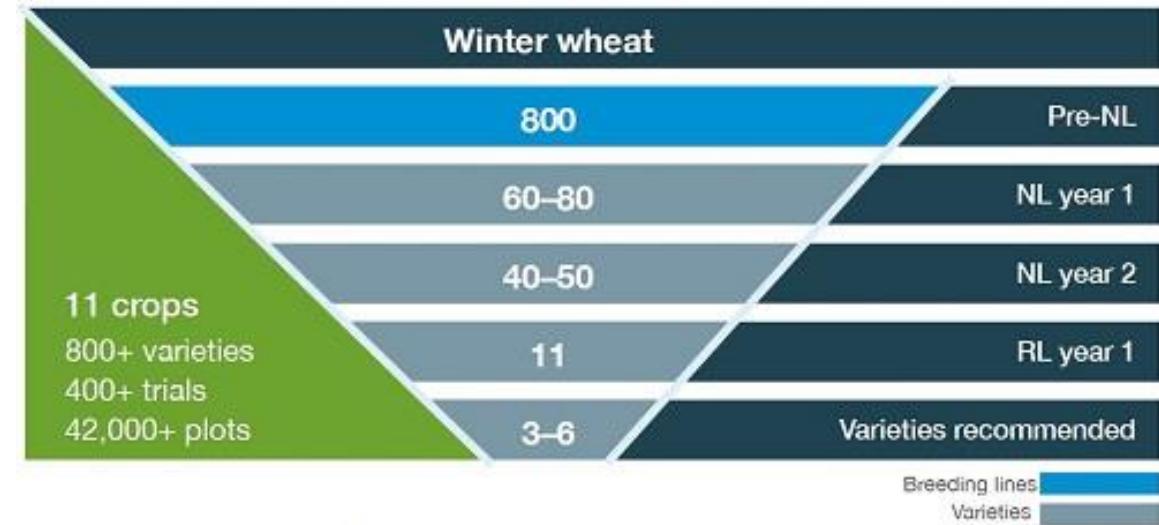
Plant Variety Registration System

Candidate variety
(UK)



Trials with fewer
varieties, more sites.

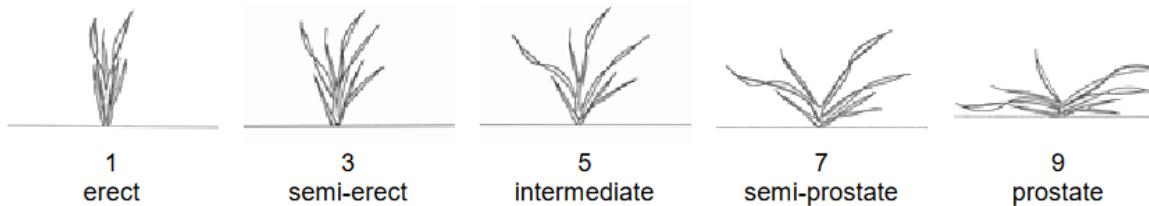
Recommended
List (RL)



DUS in barley

Ad. 2: Plant: growth habit

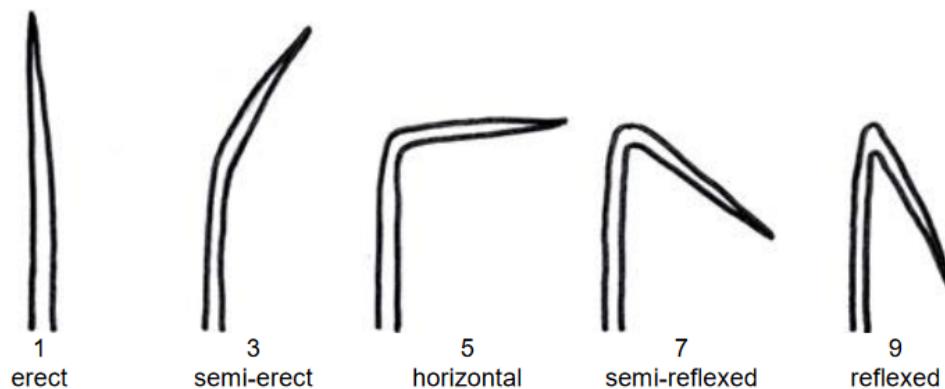
The growth habit should be assessed visually from the attitude of the leaves and tillers. The angle formed by the outer leaves and the tillers with an imaginary vertical axis should be used.



Ad. 6: Flag leaf: attitude

Flag leaf attitude is sensitive to the stage of plant development. Therefore, observation at the appropriate stage (stage 49–51 of the Zadoks decimal code) is of particular importance.

Flag leaf attitude relates to the angle between the main axis (stem) and the flag leaf blade. The expression of the majority of plants should be recorded without considering individual plants which may express a different attitude.



A set of 28 morphological traits.

Quench - Spring Barley

EAR

Ear - no. of rows : 2
Ear - glaucosity : medium
Ear - attitude : semi-erect
Ear - awn length compared to ear : longer
Ear - density : medium

PIGMENT

Pigment - awns : medium to weak
Pigment - lemma nerves : weak to medium
Pigment - flag leaf auricles : present, strong

FIRST RACHIS SEGMENT

First rachis segment - length : medium to short
First rachis segment - curvature : moderate
First rachis segment - margin hairs : numerous

STERILE SPIKELET

Sterile spikelet - attitude : divergent
Sterile spikelet - glume + awn length : medium
Sterile spikelet - palea length c.f. lemma : medium

GRAIN

Grain - rachilla hair type : short
Grain - rachilla length : medium
Grain - inner lateral nerve spicules : absent
Grain - outer lateral nerve spicules : numerous
Grain - ventral crease hairs : absent
Grain - iodicules : collar
Grain - aleurone colour : white

ADDITIONAL CHARACTER(S)

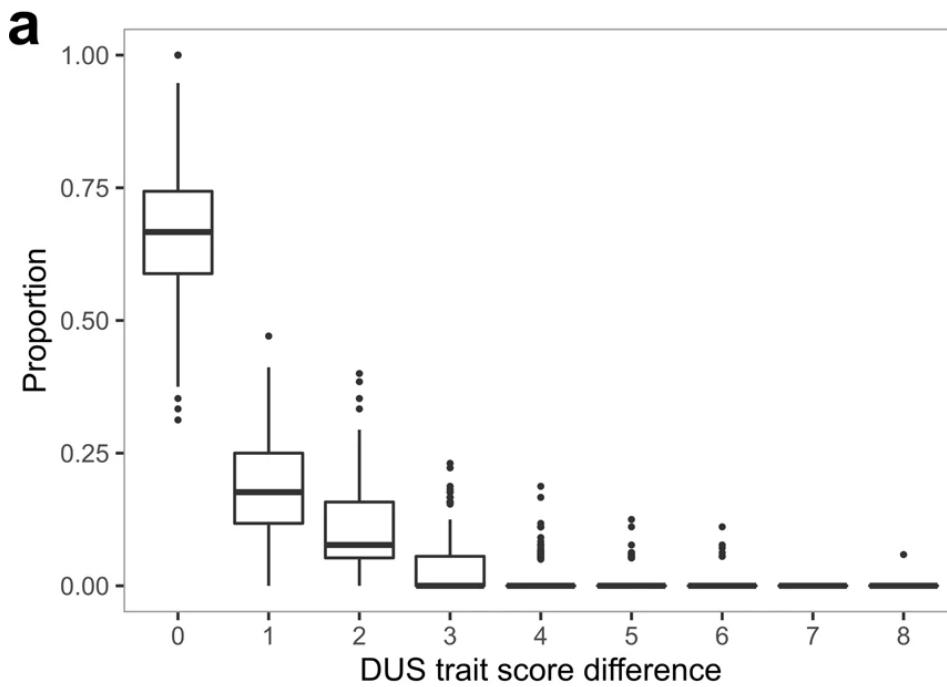
Additional character(s) - sheath hairs : absent

STERILE SPIKELET

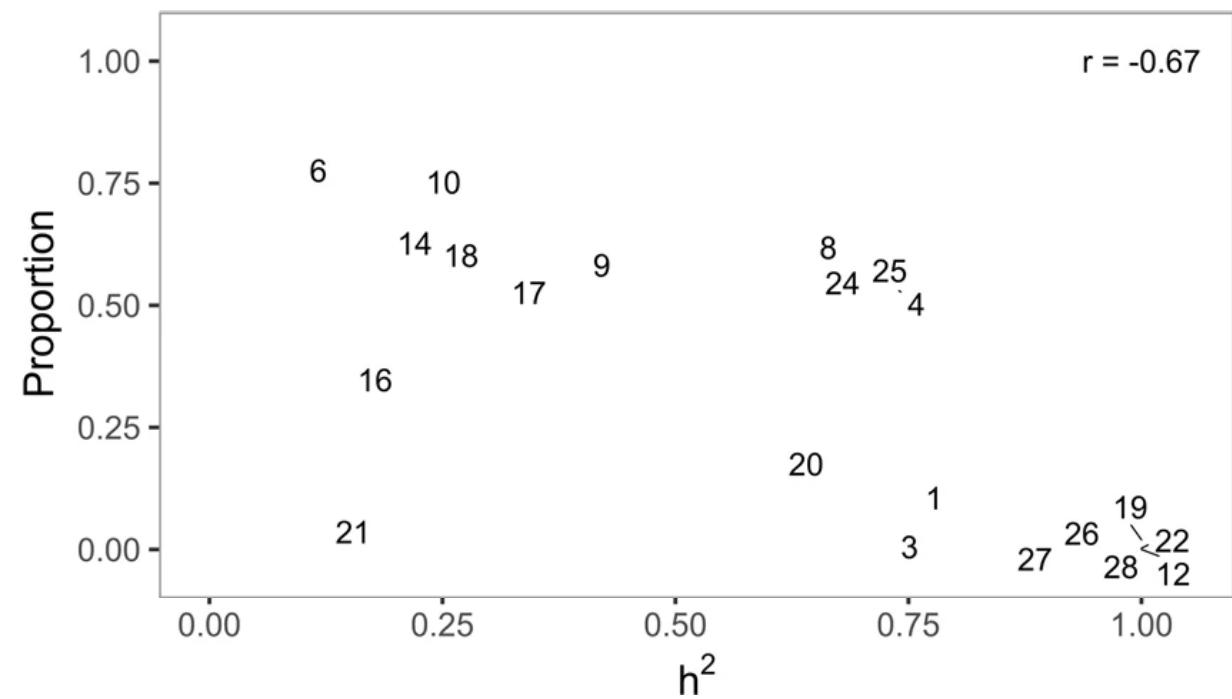
Sterile spikelet - lemma tip : rounded

DUS in barley: issues

On average, for each variety, more than a quarter of the DUS traits disagree between two scoring agencies.

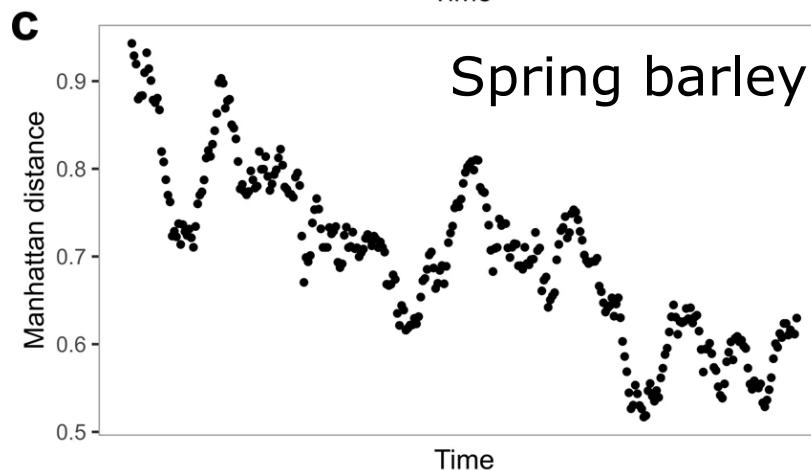
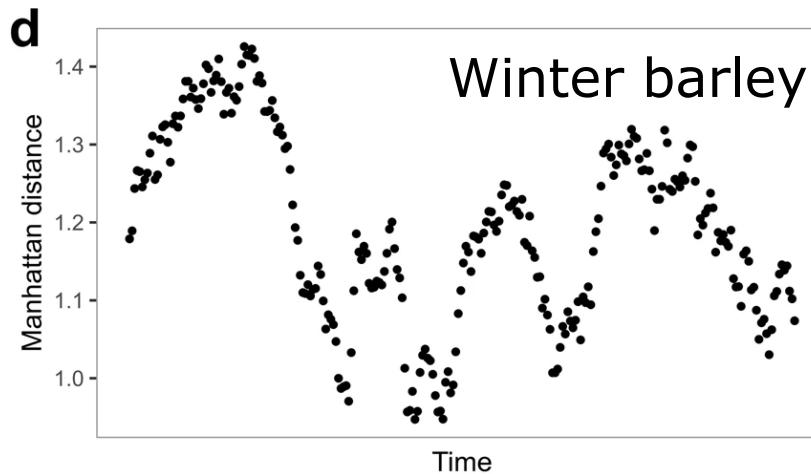


Not surprisingly, traits with high discrepancies have low heritabilities.

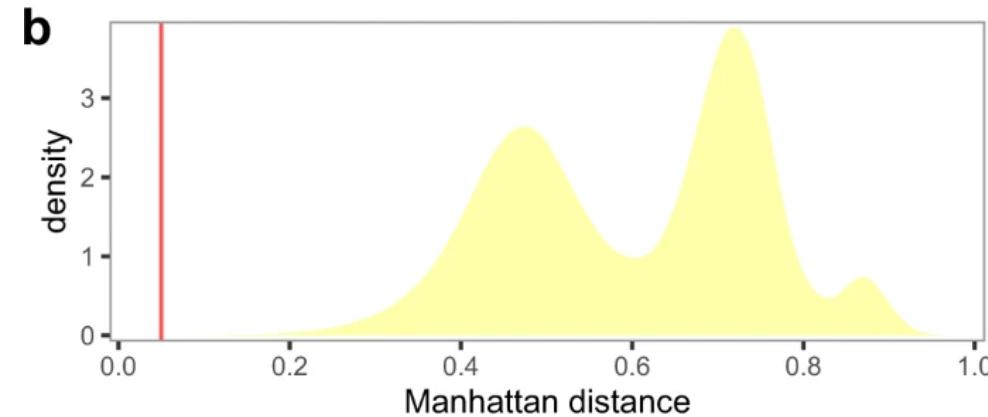


DUS in barley: issues and solutions

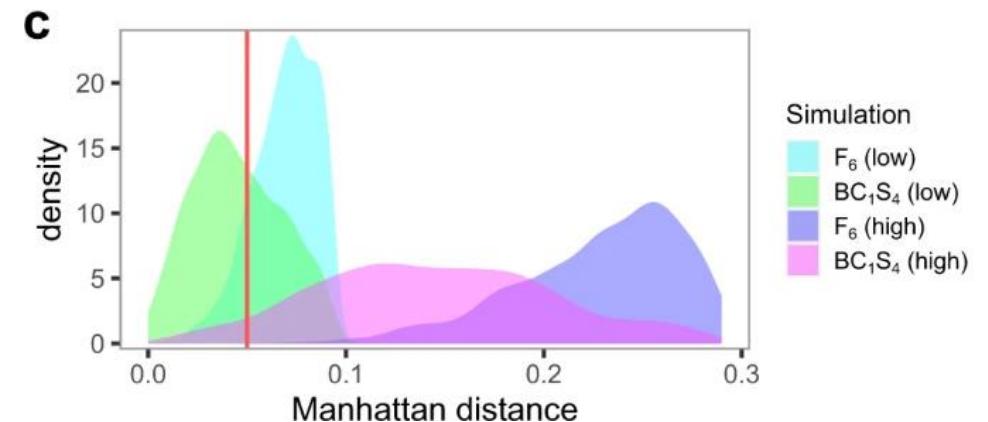
Spring varieties get more similar.



Genomic-based distance among 800 varieties



Simulated progeny from low/high distance parents



Barley Recommended List (RL)



RECOMMENDED LISTS

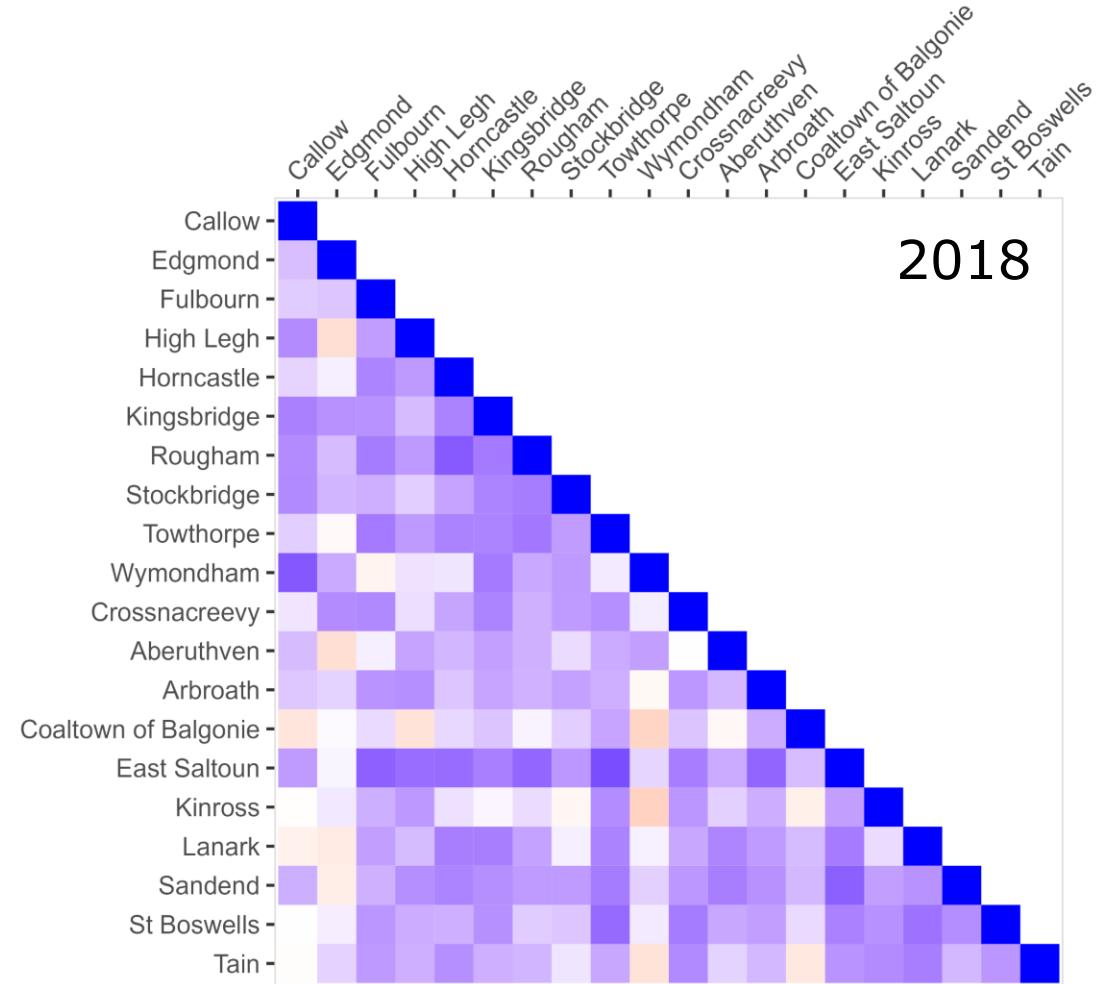
The current system does not model for G x E.

AHDB RL Harvest Results - Spring barley 2022 (Final results)

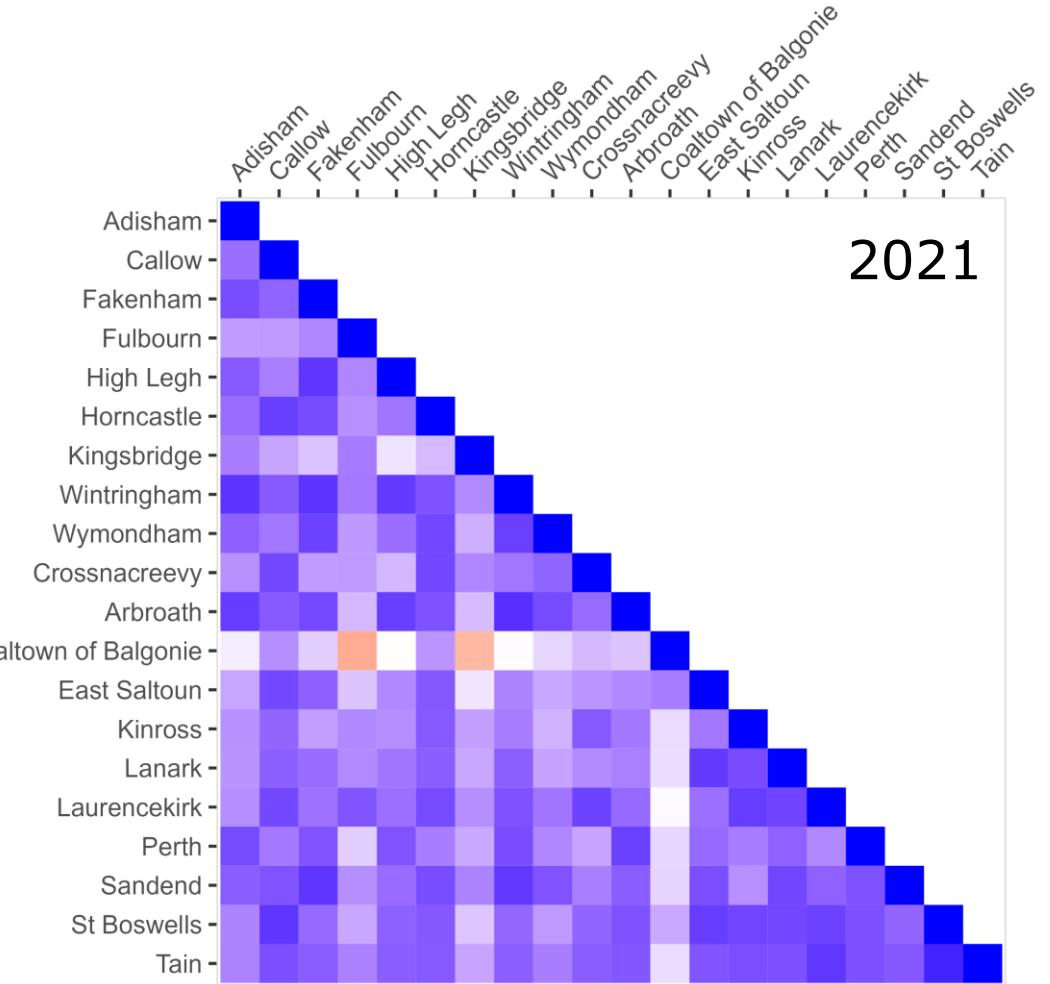
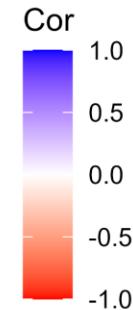
Treated grain yield (t/ha) - as % mean of controls									
07-Nov	Name	Code	18-22 mean	2022 mean	Knsgbdge Devon	Smailhlm ScBord	Callow Herefd	Fulbourn Cambs	Morley Norfk
Control mean			7.54	7.75	8.75	7.99	8.86	9.08	7.42
LSD 5% (as %C)			2.3	2.3	9.0	5.6	5.0	6.6	5.3
CV%			3.5	3.4	5.0	3.1	2.7	3.5	2.9
Sowing date				25-Mar	28-Mar	23-Mar	18-Mar	12-Mar	
Previous crop				W-wheat	S-oats	W-wheat	S-beet	S-beet	
Soil type				D-silt	Medium	Medium	Medium	Medium	
Tillage				Plough	Plough	Deep	Plough	Plough	
Propino (C)	SB2336	95	95	99	95	97	98	94	
RGT Planet (C)	SB2691	98	99	93	97	99	103	101	
Laureate (C)	SB2780	103	104	105	107	105	100	103	
LG Diablo (C)	SB2907	101	100	98	104	95	95	96	
SY Tungsten (C)	SB3068	103	103	104	98	104	104	105	
KWS Sassy	SB2771	97	98	96	99	101	100	97	
Fairing	SB2777	93	93	95	88	98	94	97	
SY Splendor	SB3073	102	102	99	101	98	103	101	
Prospect	SB3078	102	99	99	101	99	97	101	
Firefoxx	SB3085	103	103	104	105	101	104	100	

Barley Recommended List (RL)

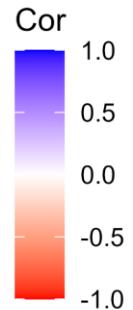
Correlations across trial sites are variable.



2018

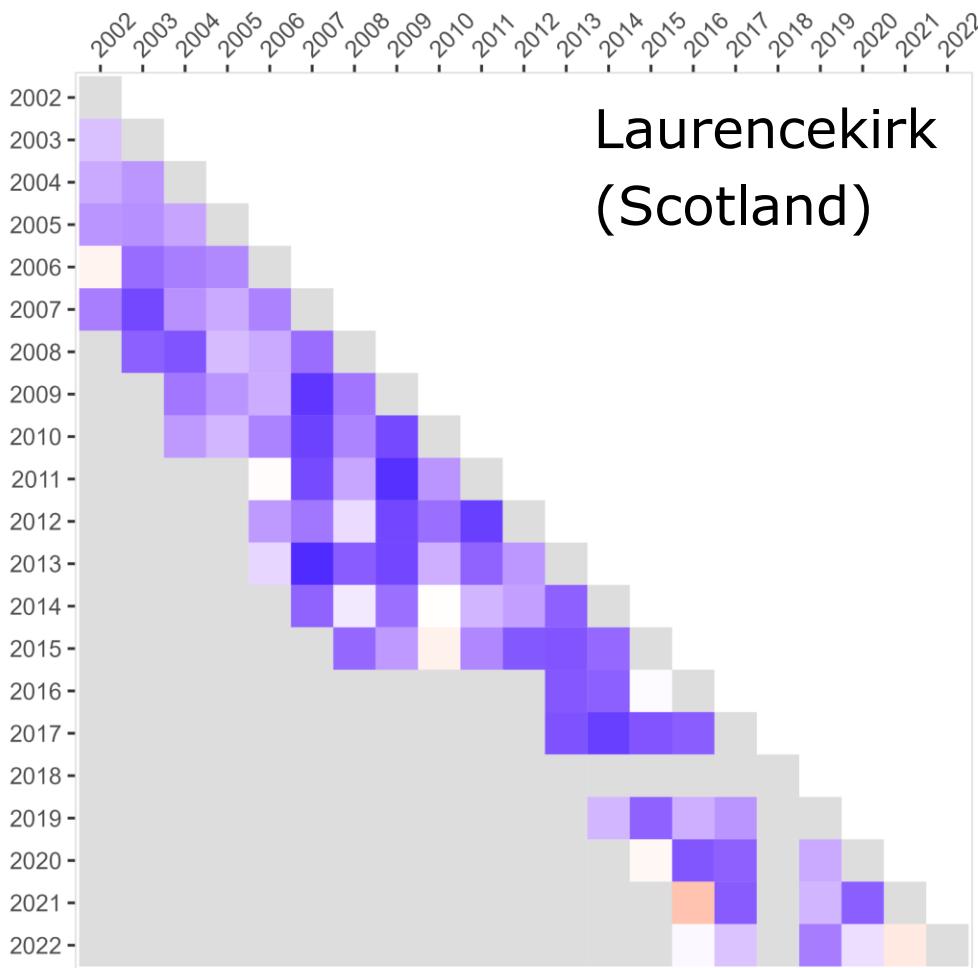


2021

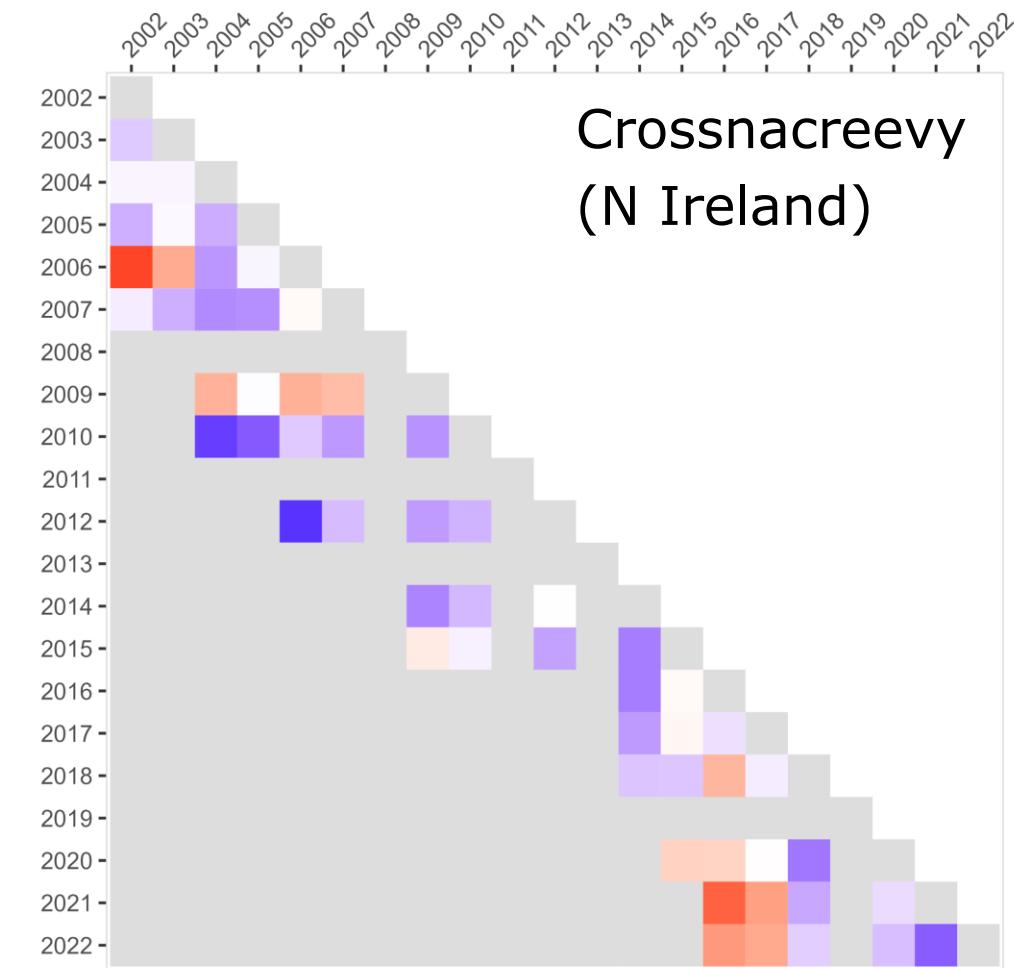


Barley Recommended List (RL)

Correlations across years are variable.

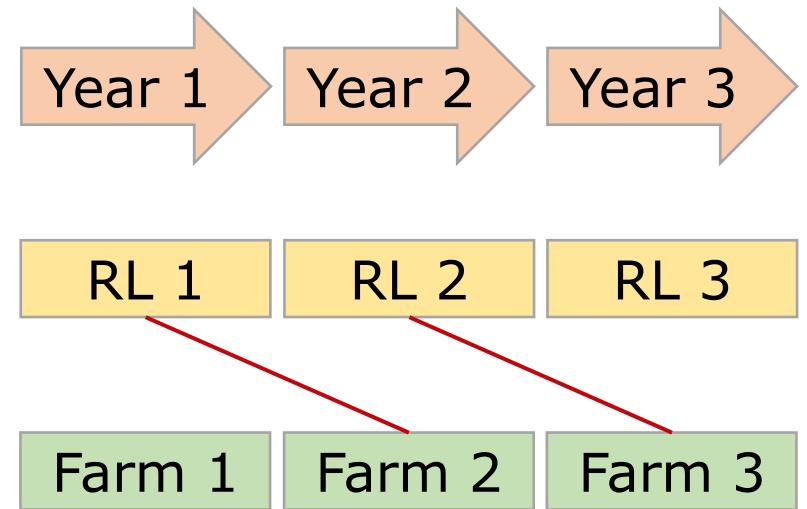
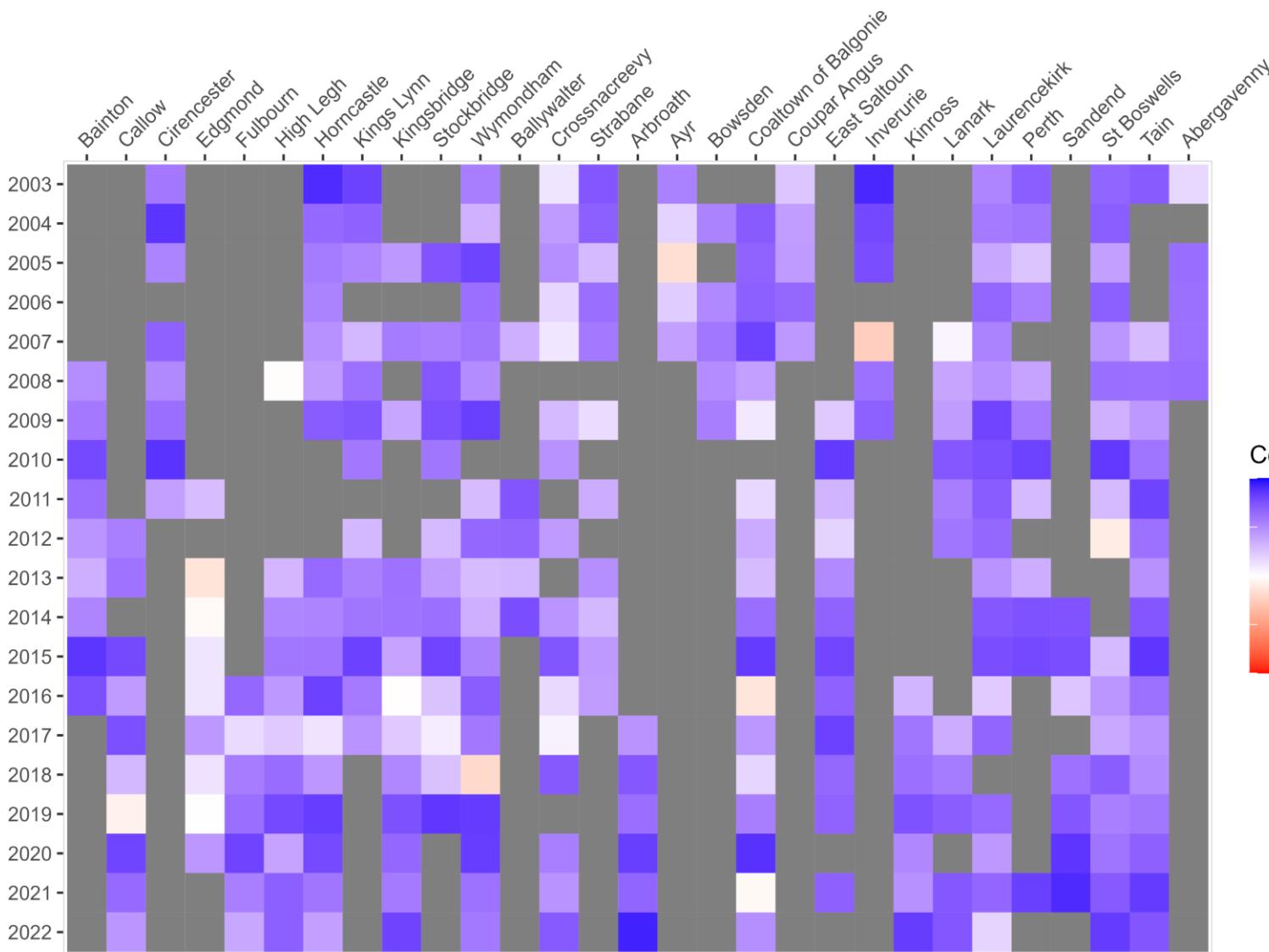


Laurencekirk
(Scotland)



Crossnacreevy
(N Ireland)

Barley Recommended List (RL)



Simple means are not
consistent predictors.

Barley Recommended List (RL)

UK RL system is way behind the time.

BLUP for phenotypic selection in plant breeding and variety testing

[H. P. Piepho](#) , [J. Möhring](#), [A. E. Melchinger](#) & [A. Büchse](#)

[Euphytica](#) **161**, 209–228 (2008) | [Cite this article](#)

We can model for G x E if we have the replicate data within each trial site.

$$Trait = Mean + Environment + Genotype + GxE + Residual$$

$$GxE \sim N(0, K_E \otimes K_G \sigma_{GxE}^2)$$

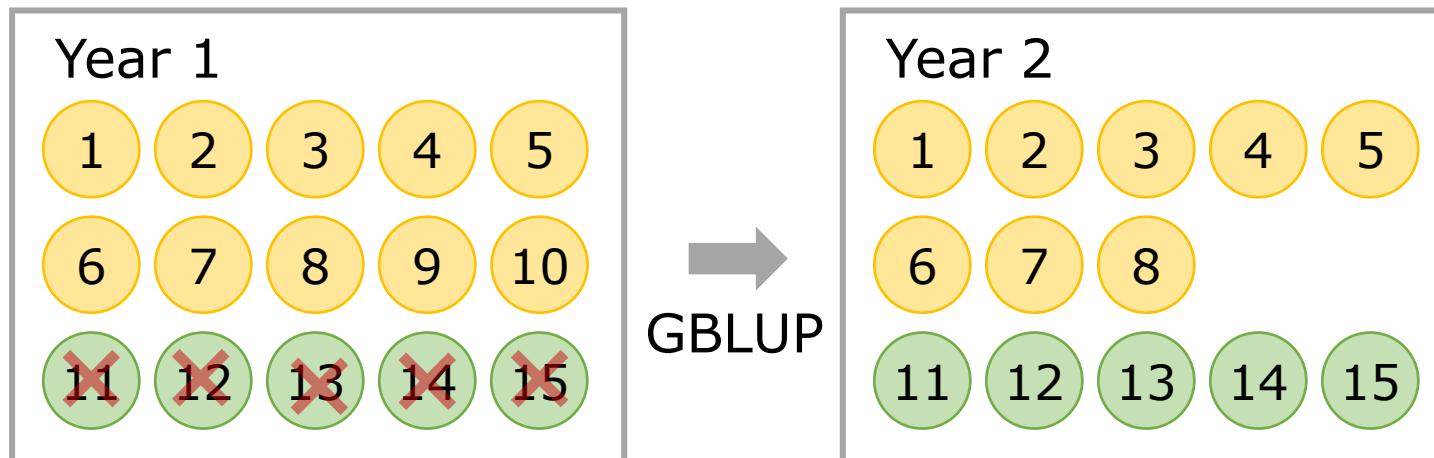
E.g. Diagonal, unstructured, factor analytic (FA).

RL system



But, we can't fit any of those models,
so what can we do?

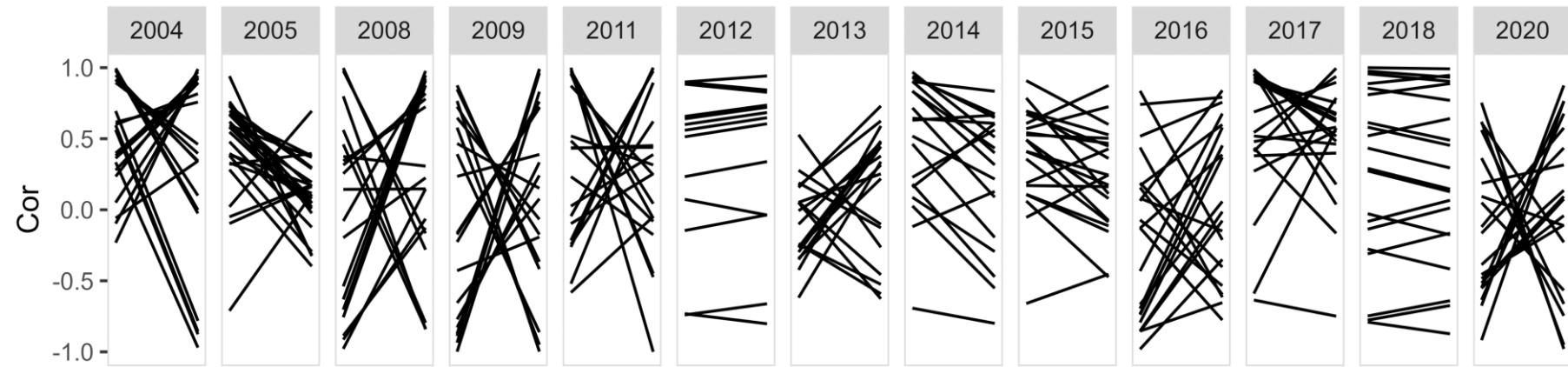
Do we really need RL?



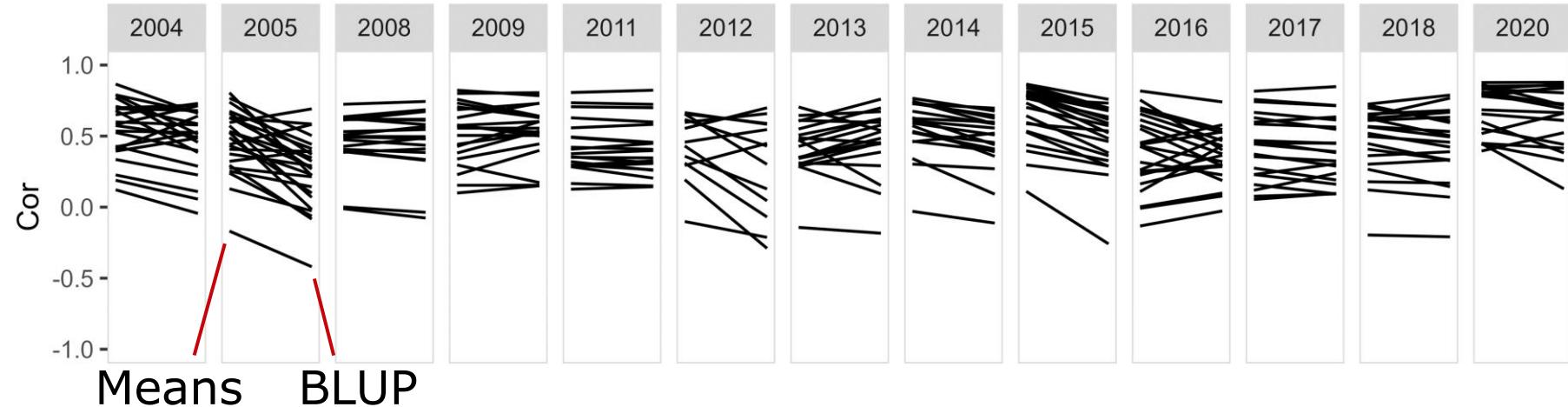
Mask the “new” varieties.

GBLUP on new varieties

New varieties only

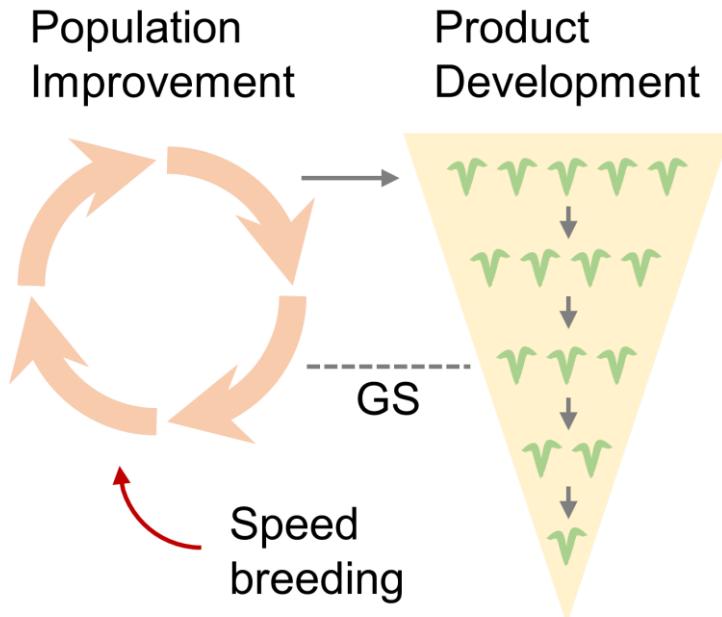


Old and new varieties

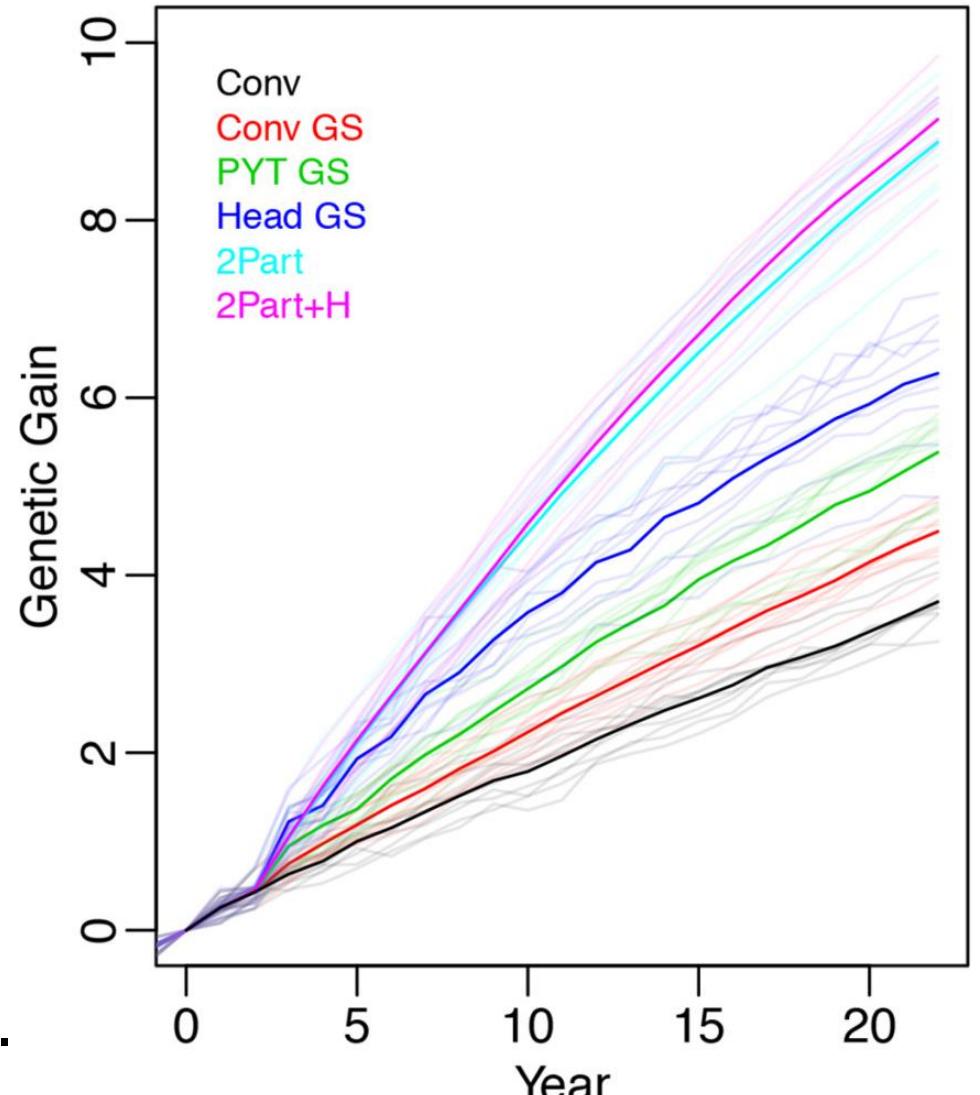


Predictions of new varieties are not the best, but it advances the varieties ahead by a year.

Breeding strategy

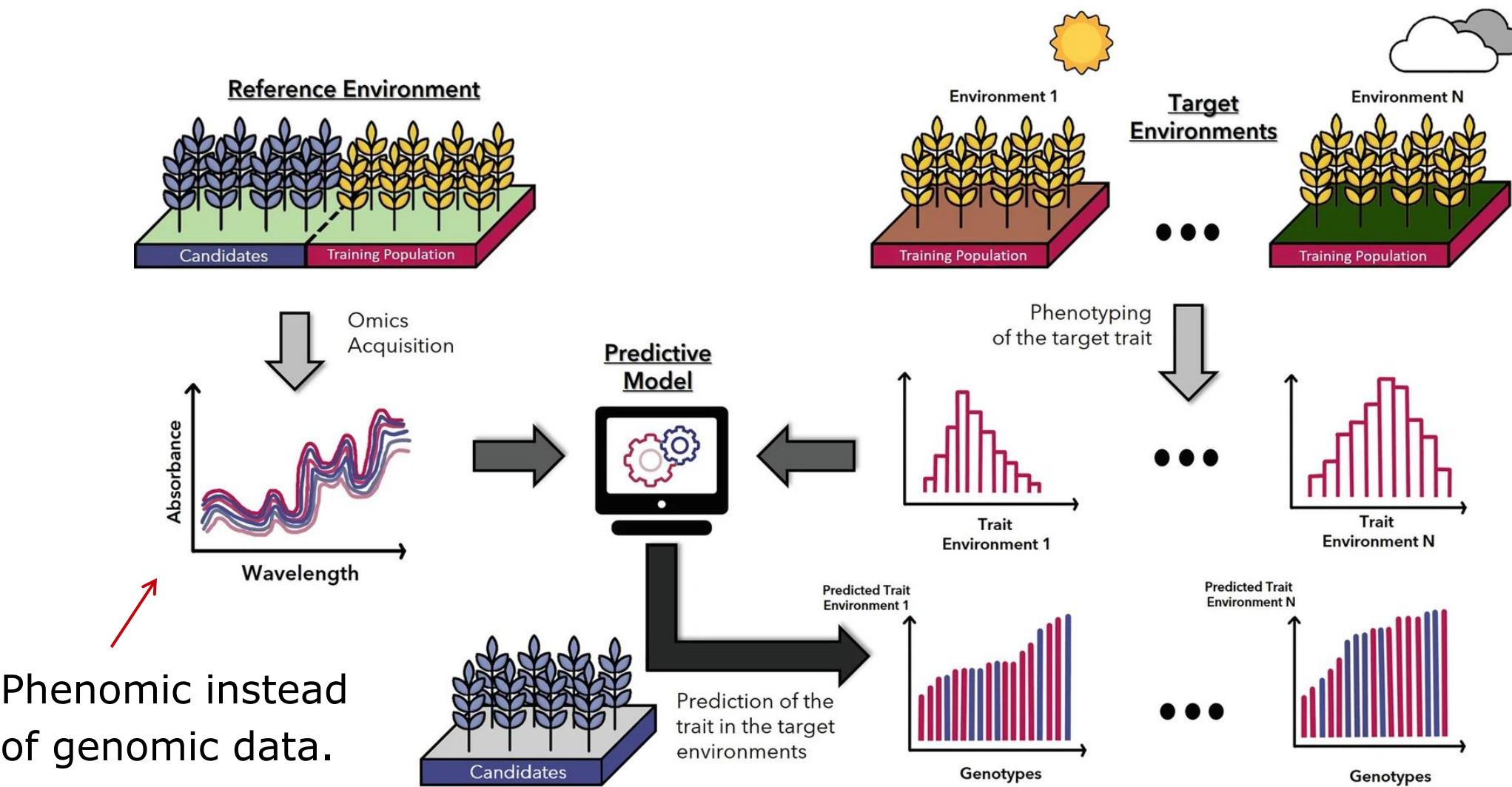


- Isolate breeding program into PI and PD.
- Allows us take advantage of GS and SB.
- Not sure about the uptake in breeding.
- Trade-off between speed and prediction accuracy.

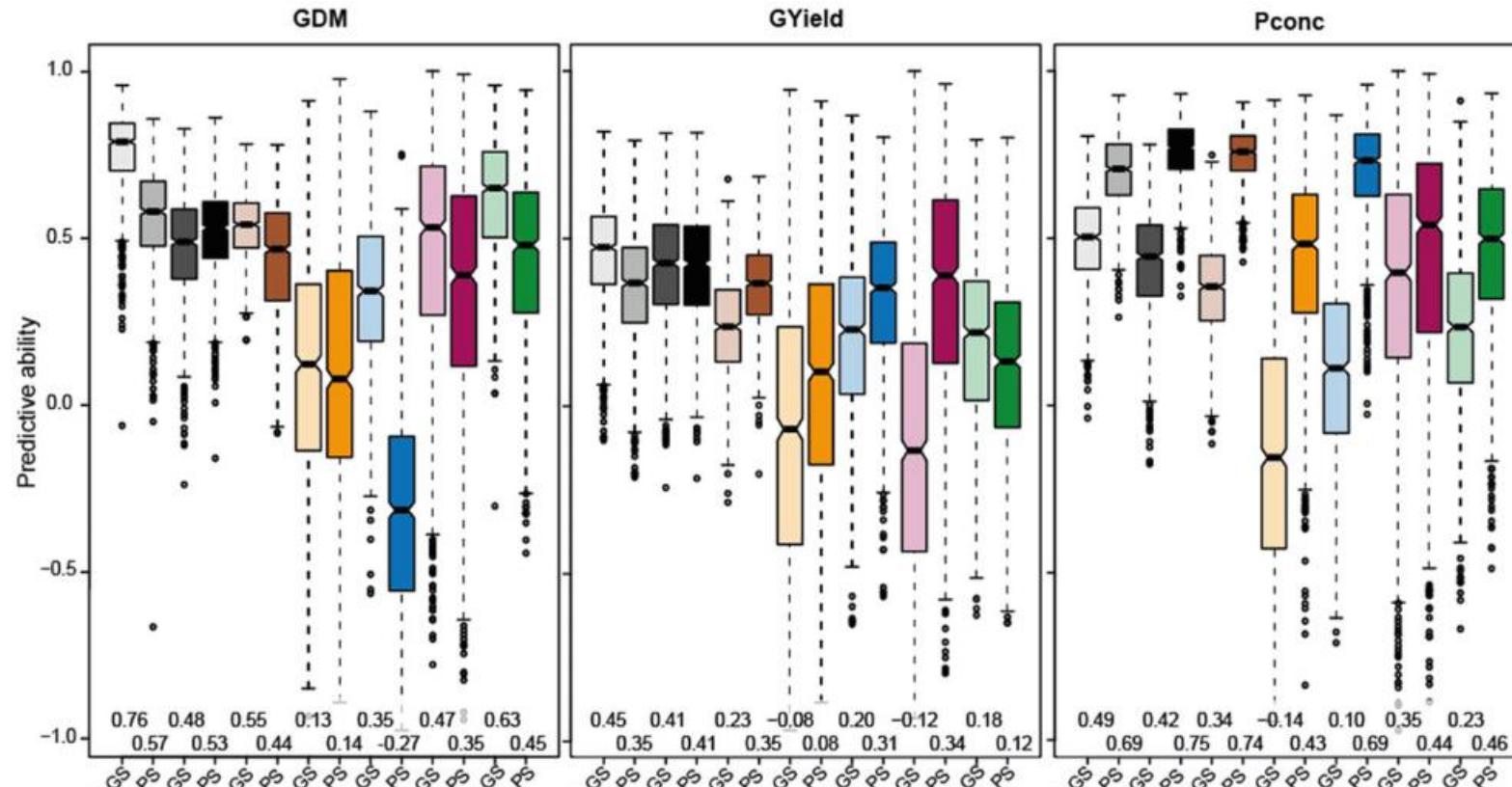


Gaynor et al (2017)

Phenomic selection

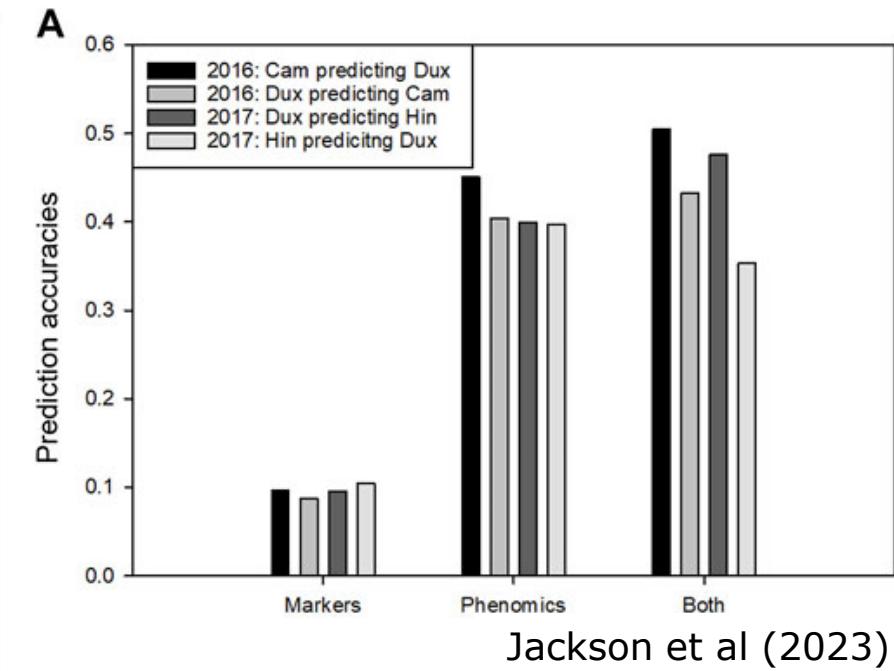


Phenomic selection



Weiss et al (2022)

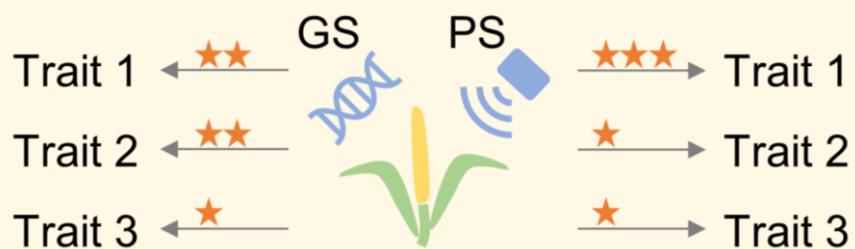
- Maize lines (elite, landrace populations).
- PS vs GS depends on populations and traits.



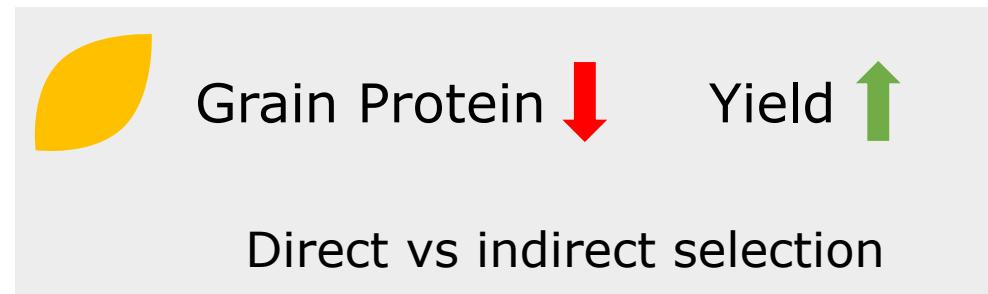
- ~40 biparental families in wheat.
- GS ignores family effect.

Room for improvement

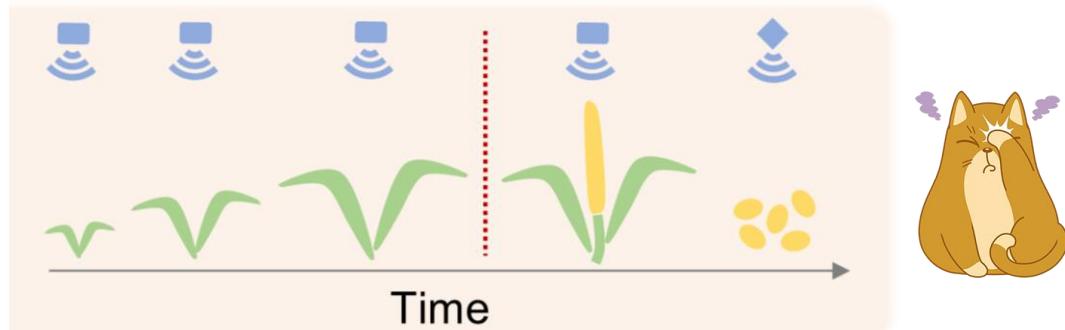
Mixed results in reported studies.



Predictions from PS are biased toward information within the predictors, Dallinger et al (2023).



Many studies use post-flowering phenotype.



Simulation is an important part of GS, how can we do that for PS?



Breeding for new environment

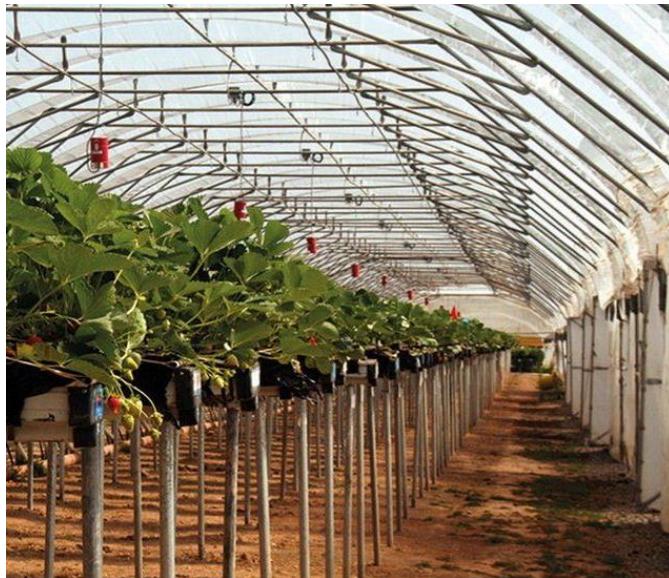
How do we breed varieties for new environment?



<https://modernfarmer.com/2015/10/can-you-grow-plants-on-mars/>

Breeding for controlled environment

Polytunnel



BBC

Aerponics



Amazon

Vertical farm



ITV

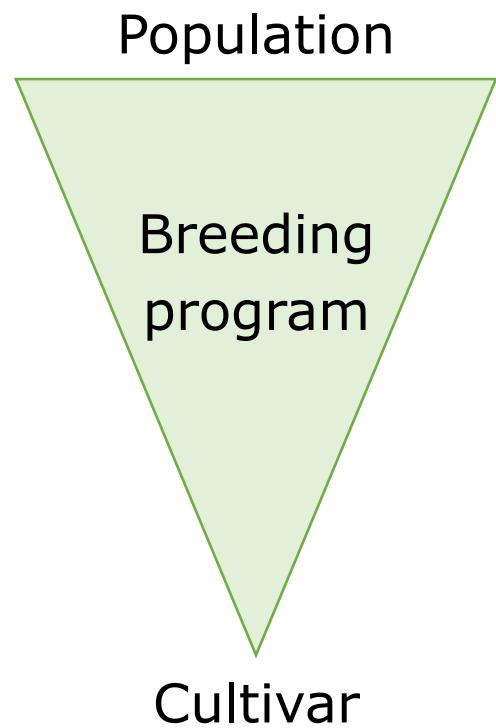
Reasons

- Climate change
- Supply chains, food miles
- Political shift

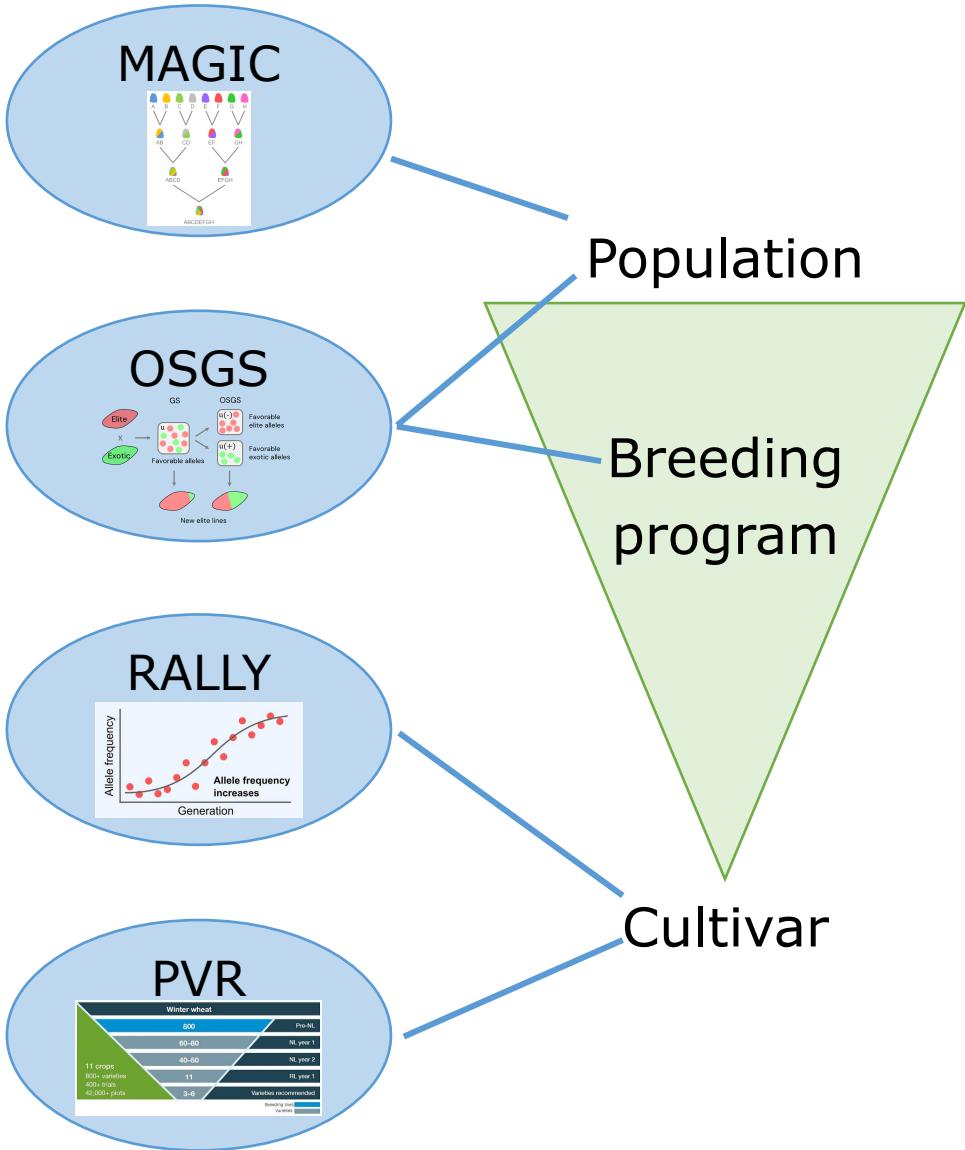
Challenges

- Energy cost
- Different disease pressure
- Lack of suitable ideotypes – need for breeding

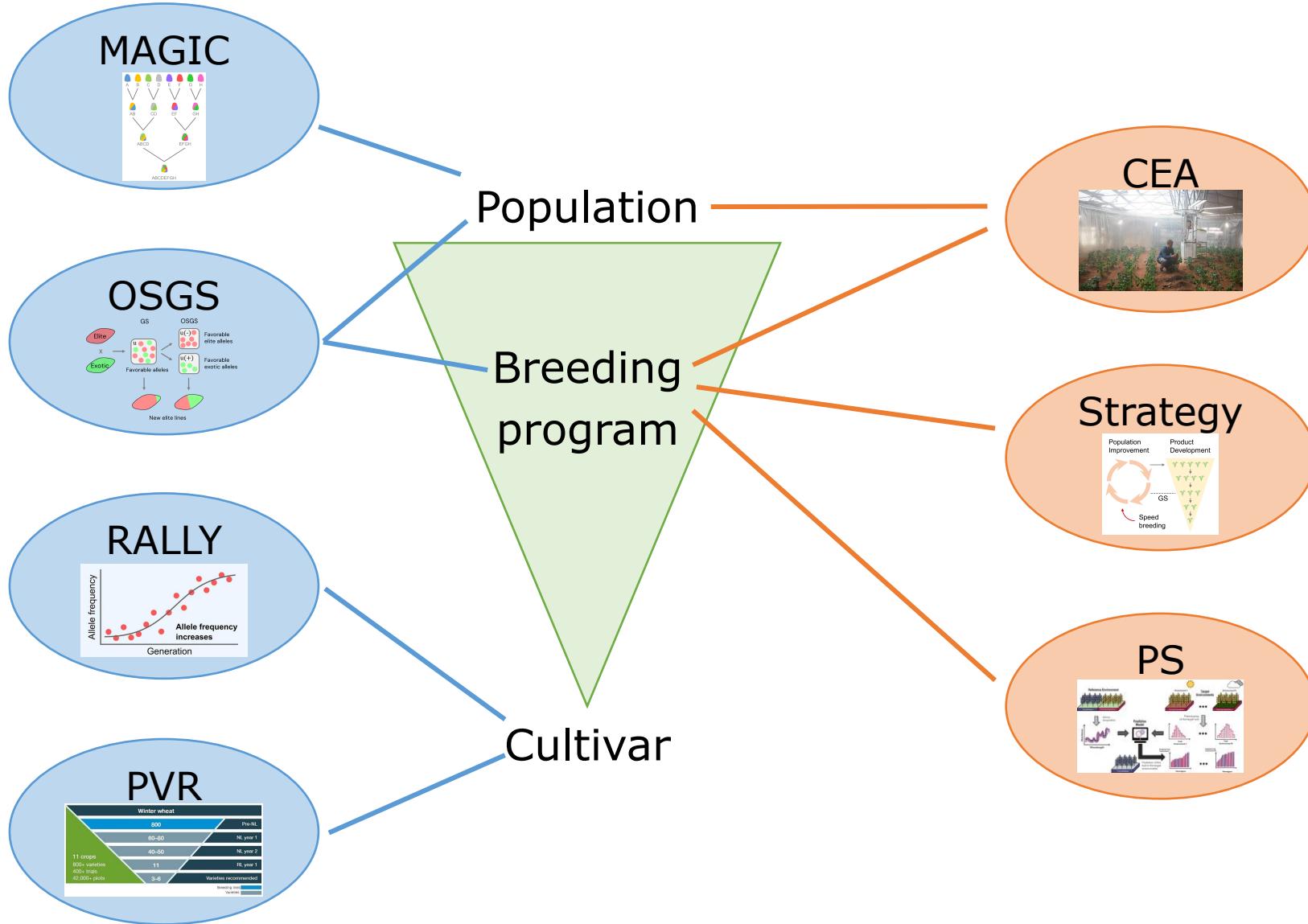
Closing statement



Closing statement



Closing statement



Acknowledgement

Organizers

Claudia Irene Calderón

Natalia de Leon

Dept of Plant and Agroecosystem Sciences

SRUC

Ian Mackay

Wayne Powell

Rajiv Sharma

Nicola Rossi

Emma Irving

David Marshall

OSGS

Gregor Gorjanc

Sarah Hearne

MAGIC

Rodney Edmondson

Hans-Peter Piepho

DUS

Joanne Russell

Luke Ramsay

Bill Thomas

RALLY

Funmi Ladejobi

Richard Mott

RL

Scottish Society for Crop Research

Computing

UK Crop Diversity



<https://cjyang-work.github.io/>



cjyang90@uwalumni.com



@hataraku_cj

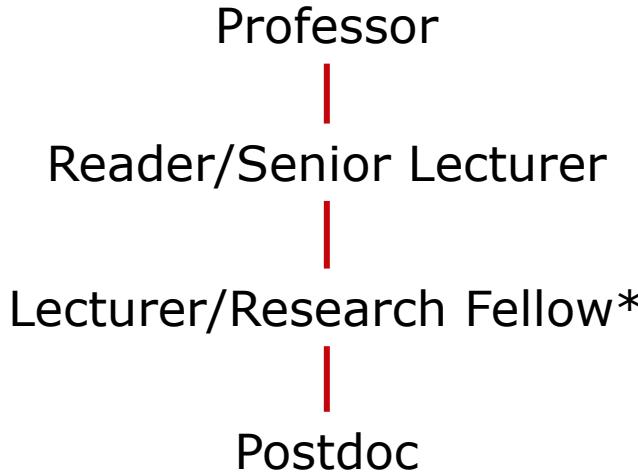


cjyang90



Research life in the UK

Academic structure



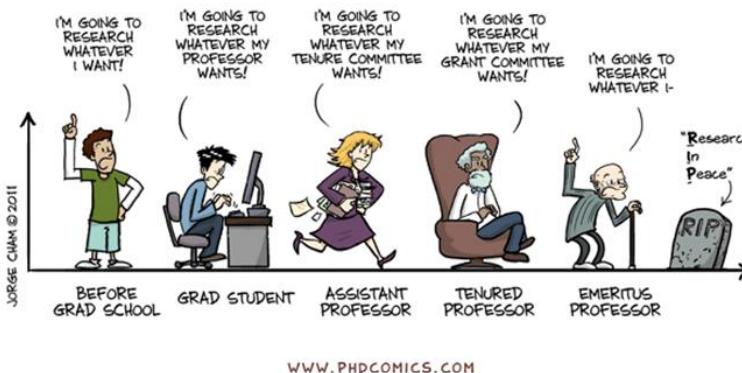
Funding body

UKRI	Others
AHRC	ERC
BBSRC	Horizon
EPSRC	Royal Society
ESRC	Innovate UK
MRC	DEFRA
NERC	RESAS
STFC	

Funding types

- Individual fellowship
 - BBSRC New Investigator
 - RS Univ. Res. Fellowship
 - Newton Int'l Fellowship
- Standard Res. Grant
- sLOLA
- Responsive Mode
- Feasibility

THE EVOLUTION OF INTELLECTUAL FREEDOM



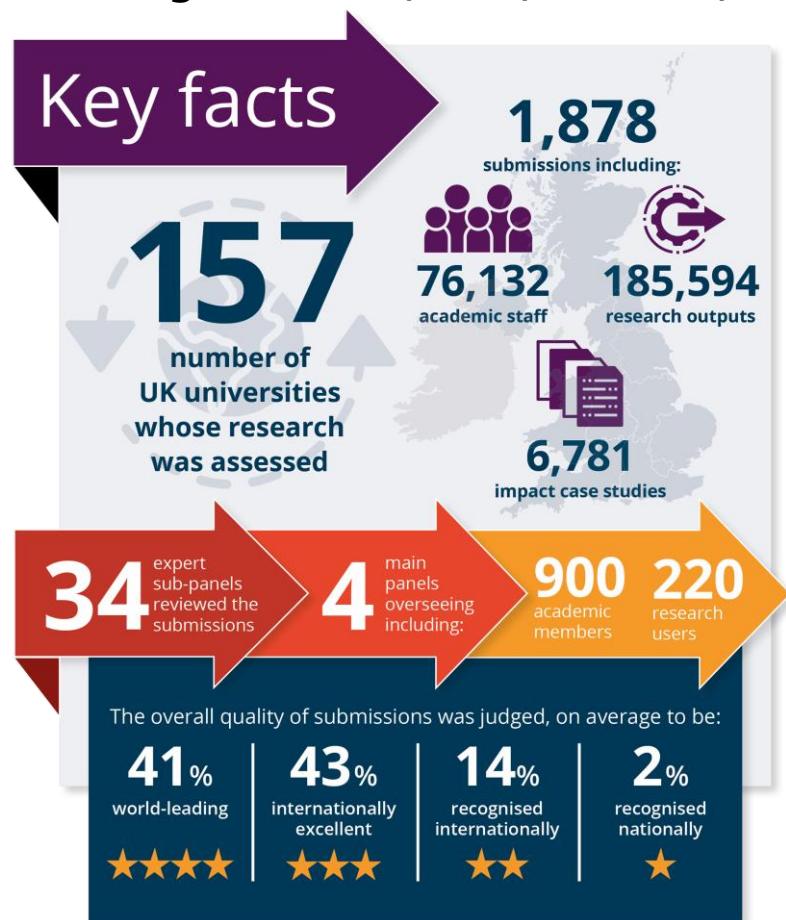
Misc.

- 2-5 years
- Work Packages (WP)
- Matched funding
- 80% fEC
- Cash vs In-kind contribution
- Workshop
- Gantt chart
- Eligibility (perm. contract)
- PI vs Co-I vs Res. Co-I
- Narrative CV
- VAT
- Host institution
- PhD (3 years, supervisors)

Research life in the UK

Research Excellence Framework (REF)

- Important for all univ.
- REF2028 is next
- Funding from RE/SFC/HEFCW/DfE



Global Talent Visa
3 years to PR

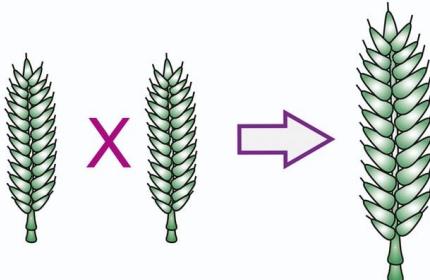


Heterosis in hybrid wheat

Heterosis = $F_1 > \text{Parents}$

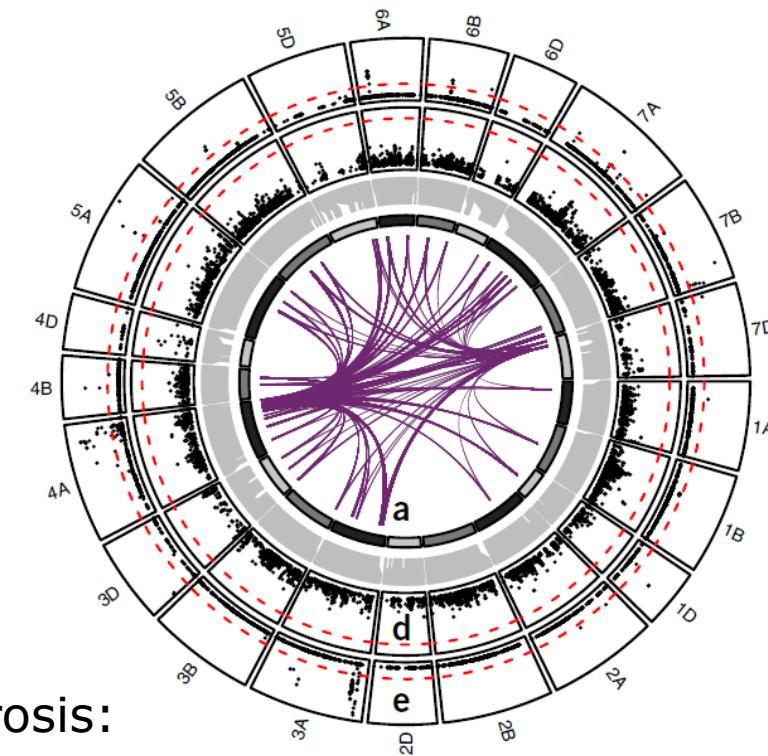
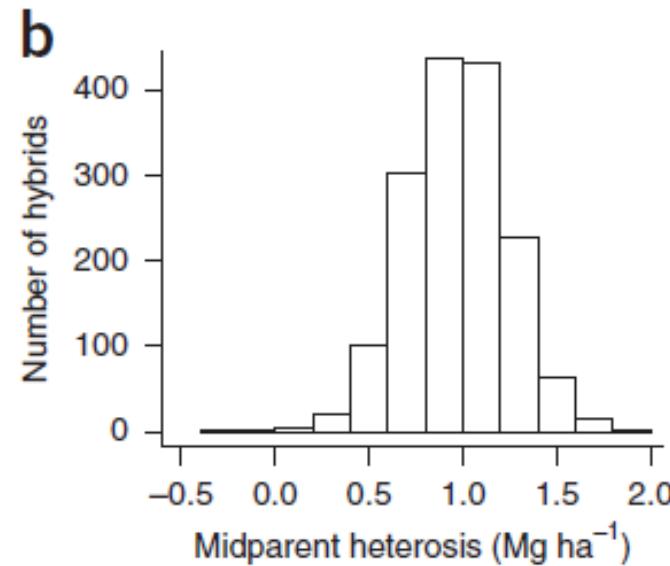
(e) Heterosis is also explained by dispersed dominant alleles

genotype: $AAbb \times aaBB \rightarrow AaBb$
phenotype: 1 x 1 → 1.5



$$AA=BB=1 \quad aa=bb=0 \quad Aa=Bb=0.75$$

Mackay et al (2020)

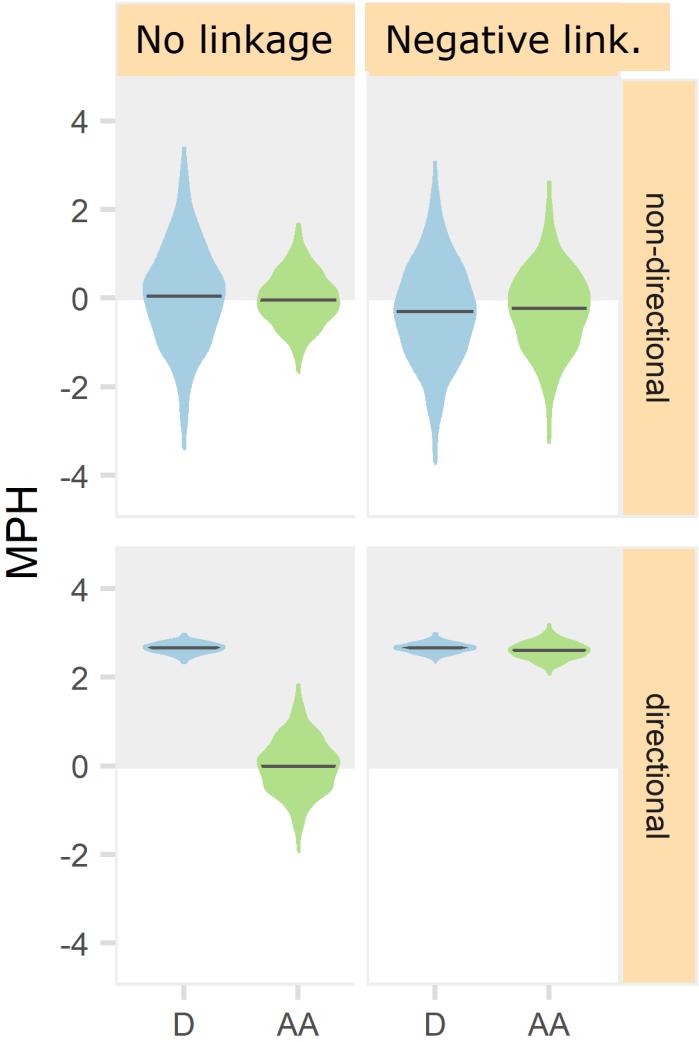


Jiang et al (2017)
Boeven et al (2020)

Genetic variance (V_G) for heterosis:

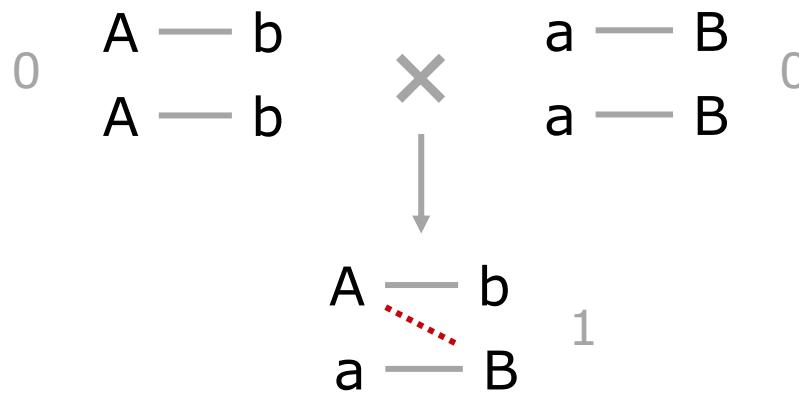
- 16/11% Dominance (V_D)
- 50/61% Additive x additive (V_{AA})
- 21/17% Additive x dominance (V_{AD})
- 13/11% Dominance x dominance (V_{DD})

Gene actions and heterosis



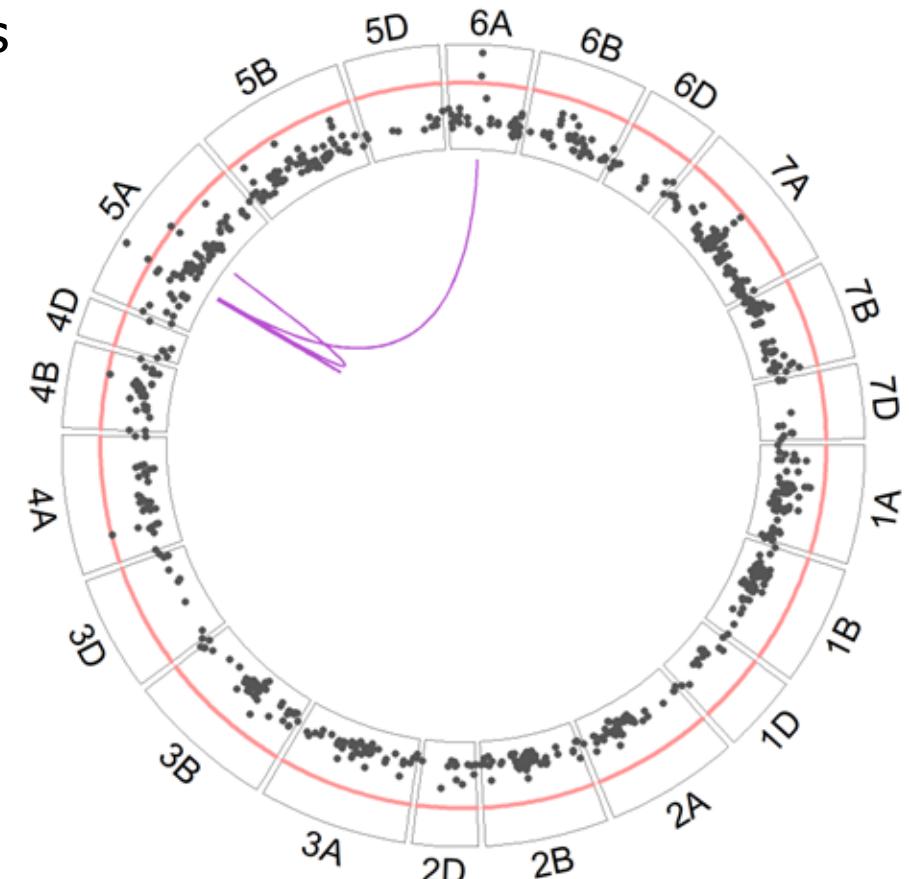
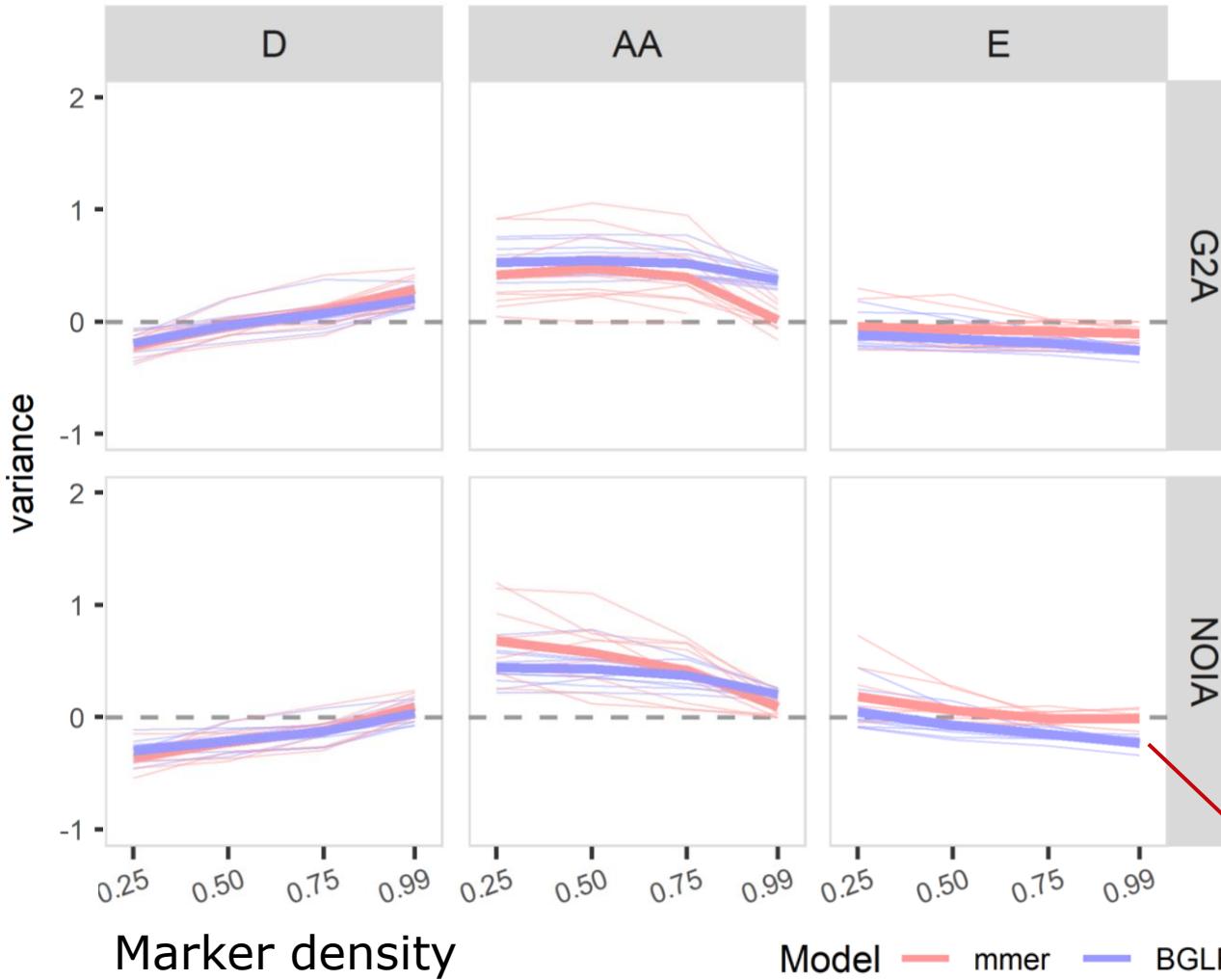
- Simulate 6 effects, 1,000 loci ($\times 2$ if linked),
100 \times 2 parents, 1,000 hybrids.
- MPH = Hybrid – Parental Average.
- Directional effect is needed for non-zero MPH.
- Linkage is needed for non-zero MPH due to AA.

Similar to “dispersed dominance” ?



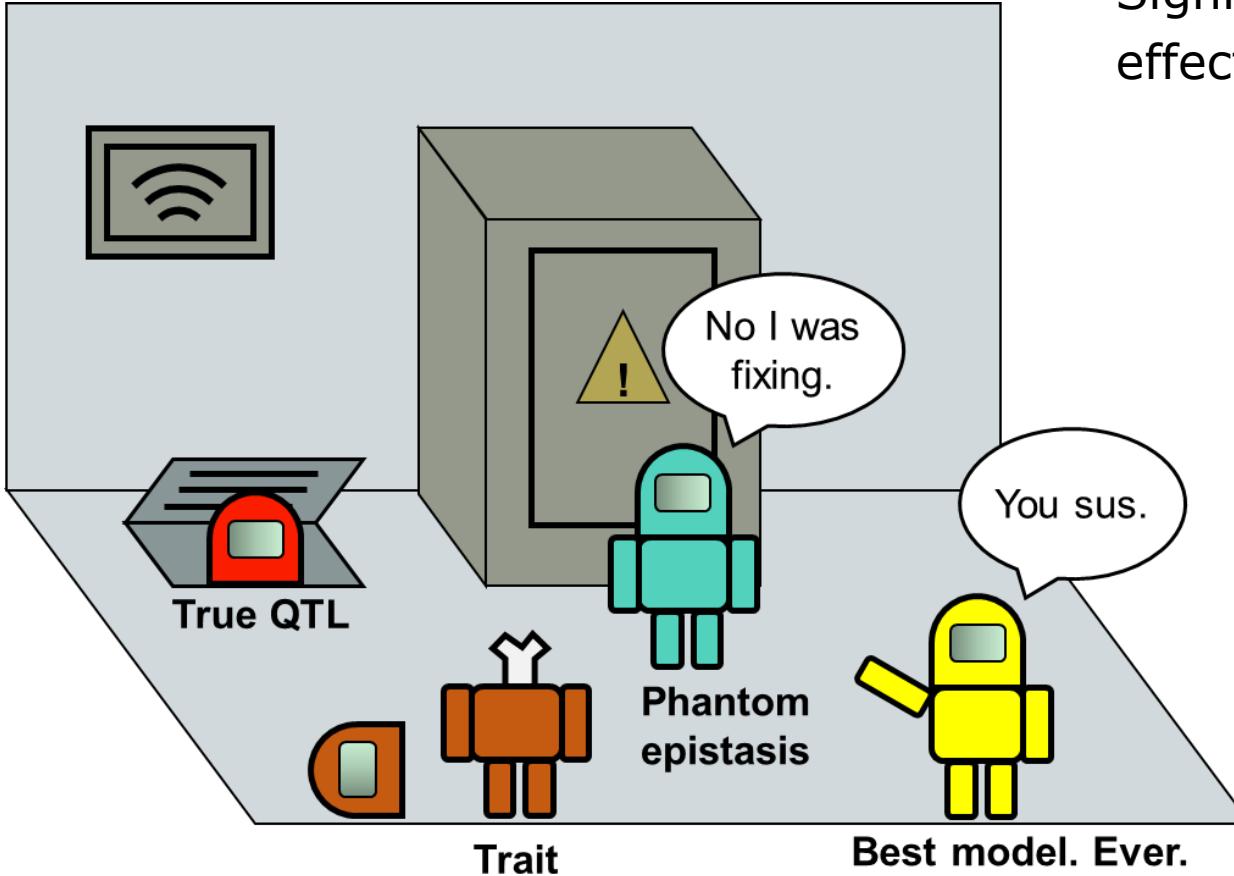
Phantom epistasis and heterosis

Simulation with purely additive and dominance QTLs



Difference between estimated
and true values.

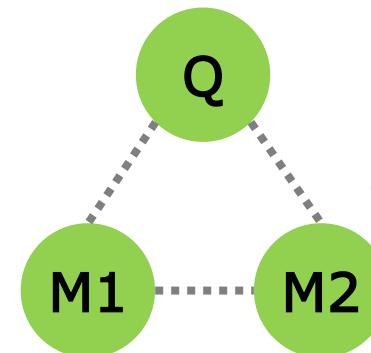
Phantom epistasis and heterosis



Phantom/apparent epistasis

Significant epistatic effect when the causative QTL effect is purely additive and/or dominance.

Imperfect linkage disequilibrium generates phantom epistasis (& perils of big data).
de los Campos et al (2019)
<https://doi.org/10.1534/g3.119.400101>



Phantom epistasis can arise between M1 and M2 if there are imperfect LDs.