Day 2
Quantitative Genetics: Dominance Effects

Chris Gaynor, Jon Bancic, Daniel Tolhurst, Gregor Gorjanc

Armidale, 2024-02-06

AlphaSimR Implementation of Dominance

- Dominance effects are a function of additive effects

$$d = \delta|a|$$

- User specifies distribution of dominance degrees

$$\delta \sim N\left(\mu_\delta, \sigma_\delta^2\right)$$

- No dominance: $\delta = 0$
- Complete dominance: $\delta = 1$
- Partial dominance: $0 < \delta < 1$
- Over-dominance: $\delta > 1$

Average Effect of an Allele Substitution

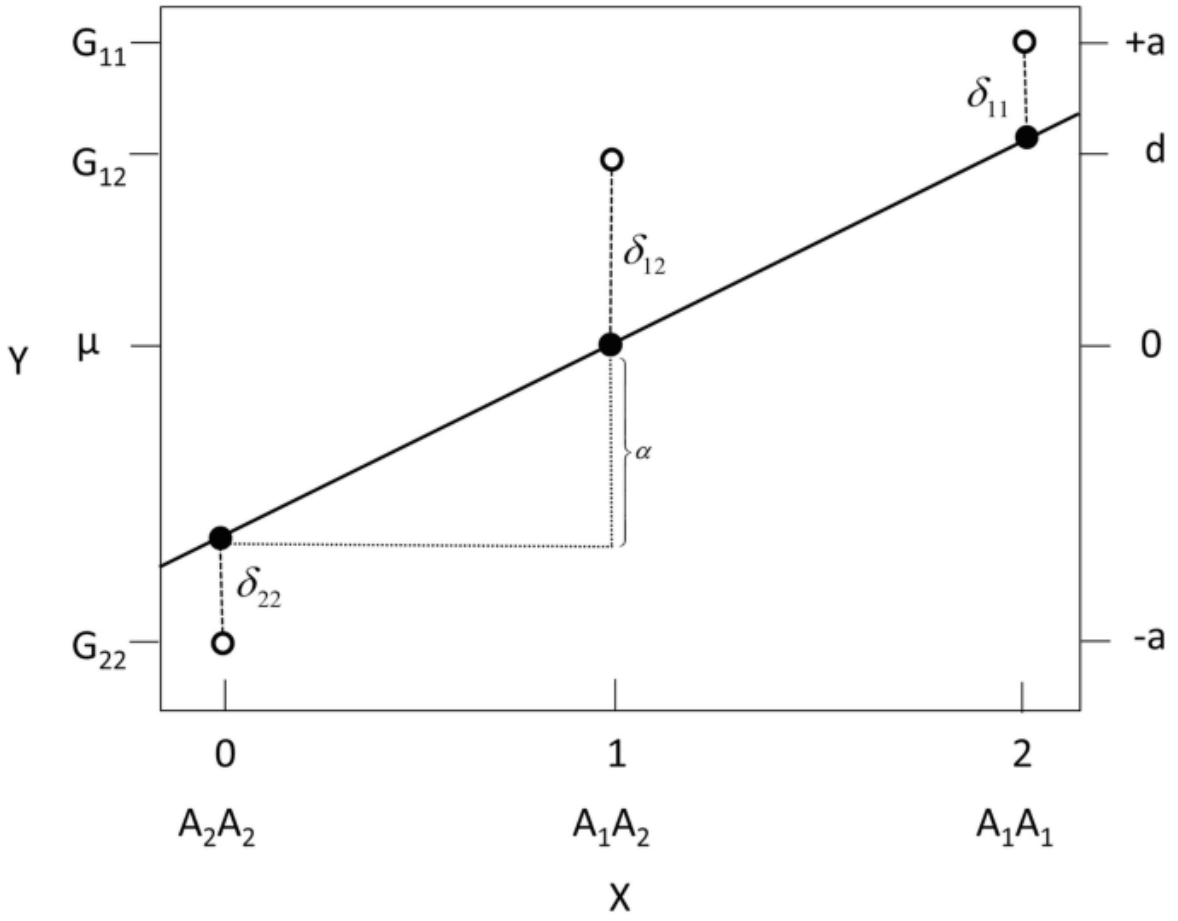
- The textbook formula assumes HWE

$$\alpha = a + (q - p)d$$

- General formula not assuming HWE

$$\alpha = a + \frac{1-F}{1+F}(q - p)d$$

- It is just a regression coefficient
  - Regress genetic value on genotype dosage
    - Assuming a base population without LD
  - Many other descriptions require HWE
    - See Falconer 1985
  - Allows for generalizations to other cases



Breeding Values

- Average effects are used to calculate breeding values
  - Breeding values are points on the regression line
- Breeding values used to calculate additive genetic variance
  - Remember this is not variance of additive effects
- There are some special properties of breeding values
  - Only valid under HWE and random mating
  - See Falconer 1985

Dominance Deviations

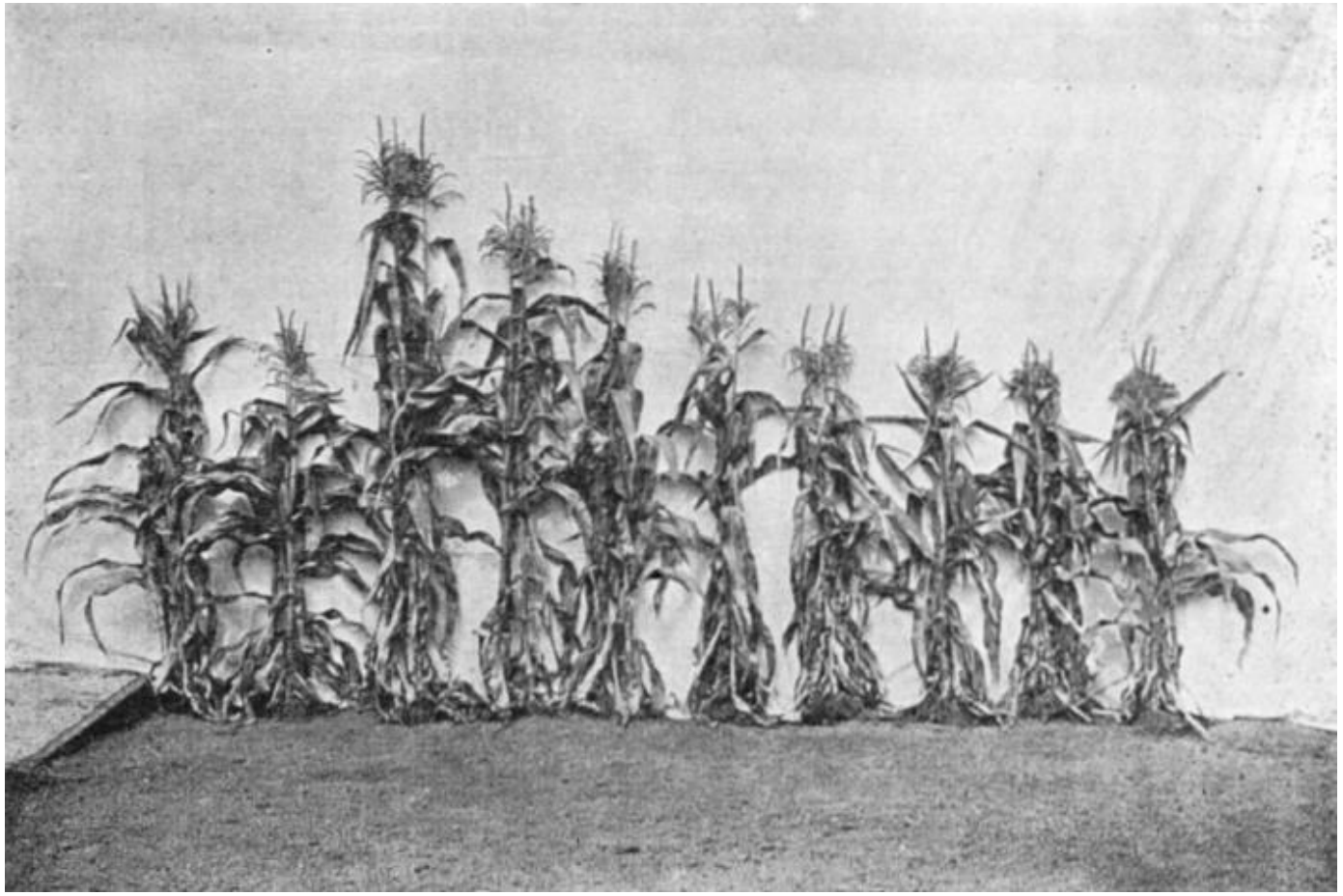
- Used to calculate dominance variance

- You can think of these as the lack-of-fit in a regression
  - Breeding values came from a regression

- You can also think of them as a quadratic polynomial
  - Constructed to be orthogonal
  - This thinking is needed for polyploids

AlphaSimR Demonstration

Heterosis and Inbreeding Depression

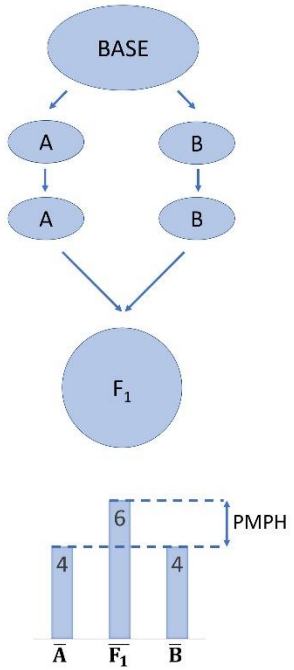


Source: Jones (1924)

Quantitative Genetics of Heterosis

Crossbreeding, Animals

Panmictic-midparent heterosis



$$\bar{F}_1 = (p - q - y)a + [2pq + y(p - q)]d$$

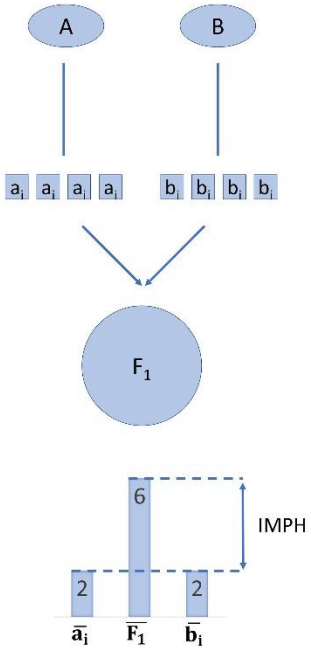
$$\overline{F_2} = (p - q - y)a + [2pq + y(p - q) - {}^1/_2 y^2]d$$

$$PMPV = (p - q - y)a + [2pq + y(p - q) - y^2]d$$

$$IMPV = (p - q - y)a$$

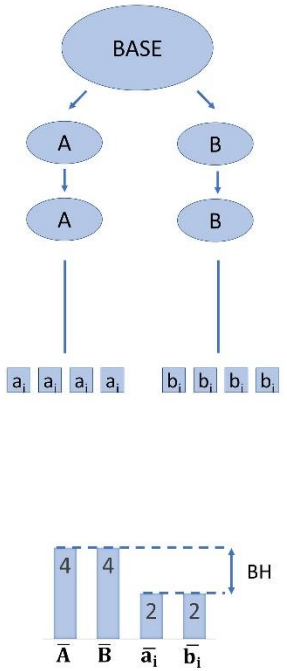
Hybrids, Plants

Inbred-midparent heterosis



Inbred vs Outbred

Baseline heterosis



$$PMPH = y^2 d$$

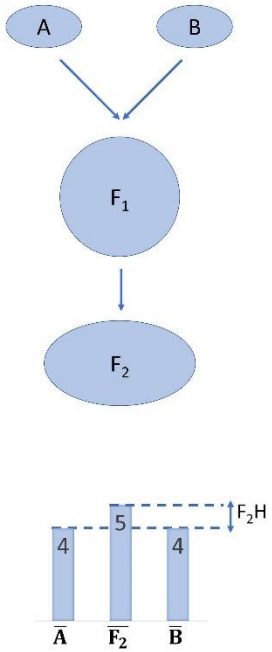
$$F_2H = \frac{1}{2}y^2 d$$

$$BH = 2\bar{p}\bar{q}d - \frac{1}{2}y^2 d$$

$$IMPH = 2\bar{p}\bar{q}d + \frac{1}{2}y^2 d$$

Inbreeding Across Lines

F_2 heterosis



Source: Labroo *et al.* (2021)

Maize Breeding Program

- Want to model genomic selection (GS) in a maize program
  - Produces hybrids and selects on general combining ability (GCA)
  - Expect GS accuracy to be lower when adding dominance
    - GCA represents a moving target

- What we know
  - Dominance gene action can explain heterosis
  - Evidence for strong partial dominance in maize
  - Measure of long-term genetic gain for inbreds and hybrids

Dominance Theories for Heterosis

Genotype	0	1	2
Genetic Value	-a	d	a

- Overdominance
  - d>a
- (Partial) Dominance
  - a>d>0
  - Pseudo-overdominance

Pseudo-overdominance Example

Locus 2 (a=1, d=0.1)	**Locus 1 (a=1, d=0.1)**		
	0	1	2
0	-2a (-2)	d-a (-0.9)	0 (0)
1	d-a (-0.9)	2d (0.2)	a+d (1.1)
2	0 (0)	a+d (1.1)	2a (2)

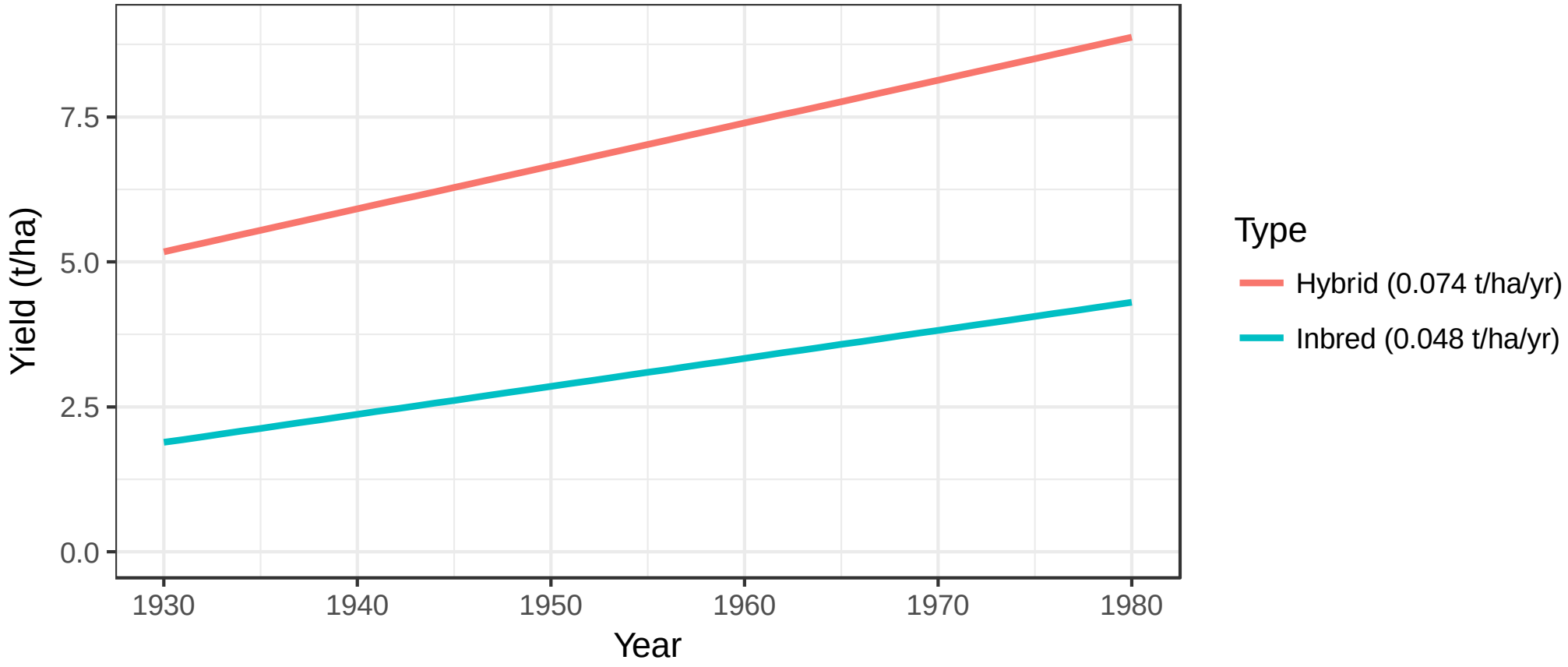
- All possible combinations
- Homozygote is best

Pseudo-overdominance Example

Locus 2 (a=1, d=0.1)	**Locus 1 (a=1, d=0.1)**		
	0	1	2
0	-2a (-2)	d-a (-0.9)	0 (0)
1	d-a (-0.9)	2d (0.2)	a+d (1.1)
2	0 (0)	a+d (1.1)	2a (2)

- Repulsion linkage (complete)
  - Unobserved combinations
- Heterozygote is best

Long-term Genetic Gain in Maize



Source: Troyer and Wellin (2009)

Quantitative Genetics for Heterosis (Hybrids)

- Inbred midparent heterosis
  - Deviation of hybrids from midparent value
  - Heterosis from two sources
    - Recovery of inbreeding
    - Split between pools

$$H_{IMP} = \sum_{i=1}^{n_{QTL}} 2\bar{p}_i\bar{q}_i d_i + \frac{1}{2}(p_{1i} - p_{2i})^2 d_i$$

Quantitative Genetics in AlphaSimR

$$a \sim N\left(0, s_a^2\right)$$

$$d = d|a|$$

$$d \sim N\left(m_d, s_d^2\right)$$

$$E\left(H_{IMP}\right) = \frac{m_d s_G \sum_{i=1}^{n_{QTL}} 2\bar{p}_i\bar{q}_i + \frac{1}{2}\left(p_{1i} - p_{2i}\right)^2}{\sqrt{2p\sum_{i=1}^{n_{QTL}} \bar{p}_i\bar{q}_i}}$$

Quantitative Genetics in AlphaSimR (cont'd)

$$E\left(H_{IMP}\right) = \frac{m_d s_G \sum_{i=1}^{n_{QTL}} 2\bar{p}_i\bar{q}_i + \frac{1}{2}\left(p_{1i} - p_{2i}\right)^2}{\sqrt{2p\sum_{i=1}^{n_{QTL}} \bar{p}_i\bar{q}_i}}$$

$$H_{IMP}\left(s_G\right) \mu s_G$$

$$H_{IMP}\left(d\right) \mu m_d$$

$$H_{IMP}\left(n_{QTL}\right) \mu y$$

$$y \gg \sqrt{n_{QTL}}$$

Setting Dominance Level for Maize

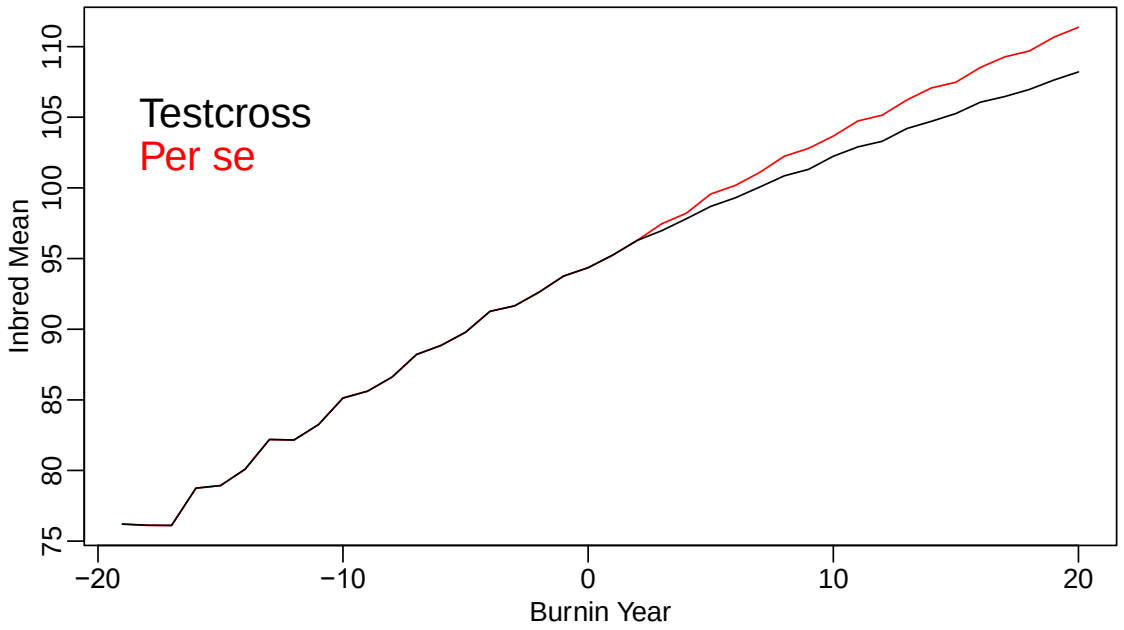
- Trade-off between two key values
  - Number of QTL
  - Mean dominance degree
  - Others matter to a lesser degree
- Estimated optimal values
  - 300 QTL per chromosome
  - 0.92 mean dominance degree
- Provides a good approximation to long-term genetic gain
  - Reasonable but not necessarily realistic

About that Troyer and Wellin Paper

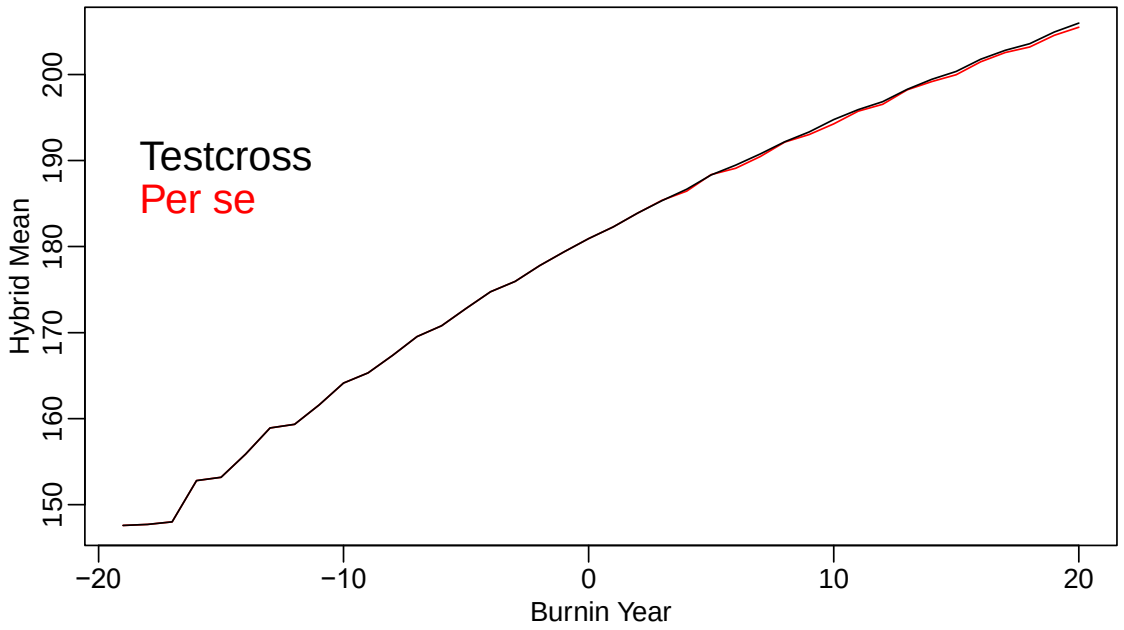
- They argued for eliminating the first round of testcrosses
  - To be replaced with a *per se* evaluation of inbreds
- We can examine this in simulation
  - Using the dominance values from before
- Simulation supports their recommendation

Replacing First Testcross with *per se* Evaluation

Inbred genetic gain



Hybrid genetic gain



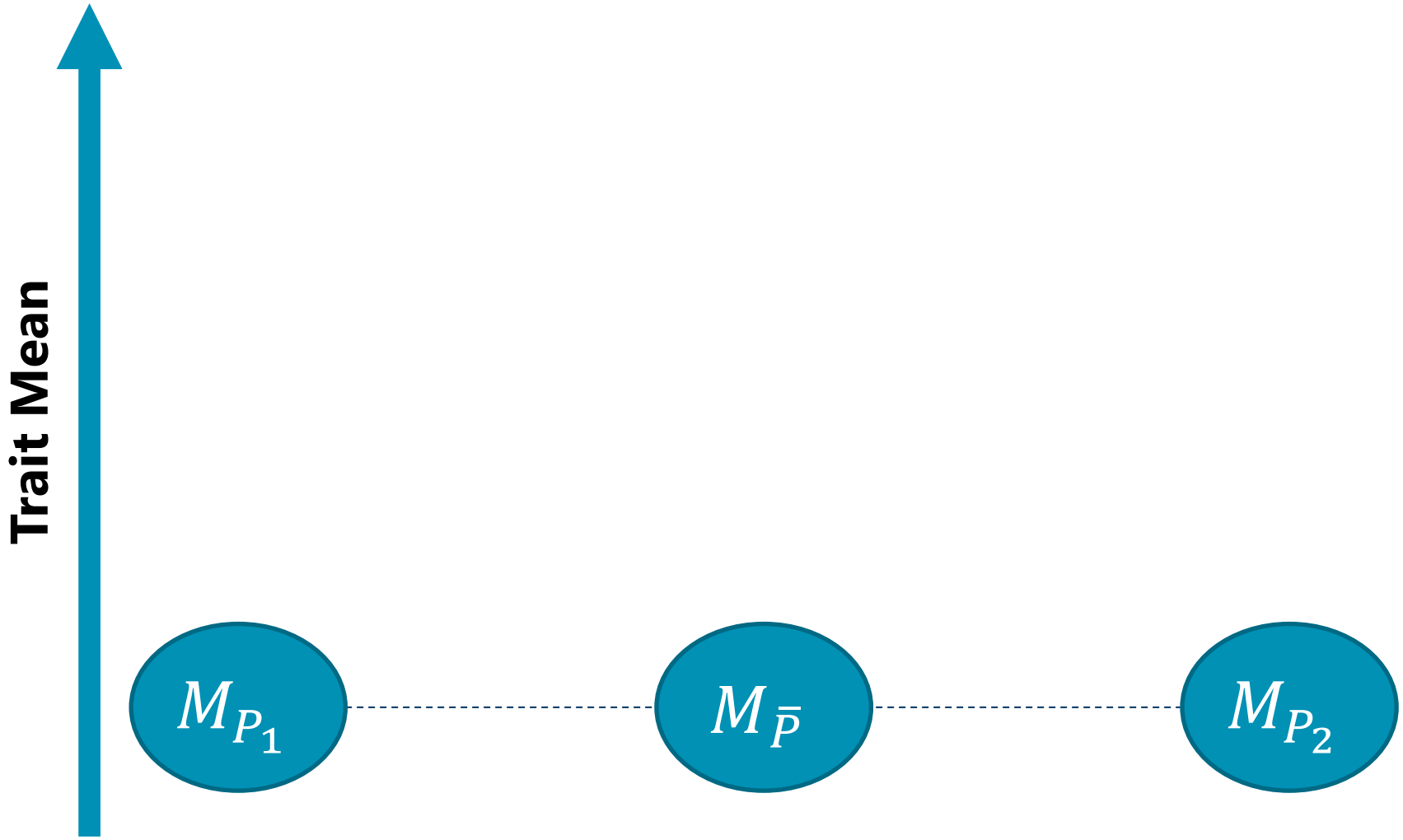
AlphaSimR Demonstration

Heterosis in Animal Breeding

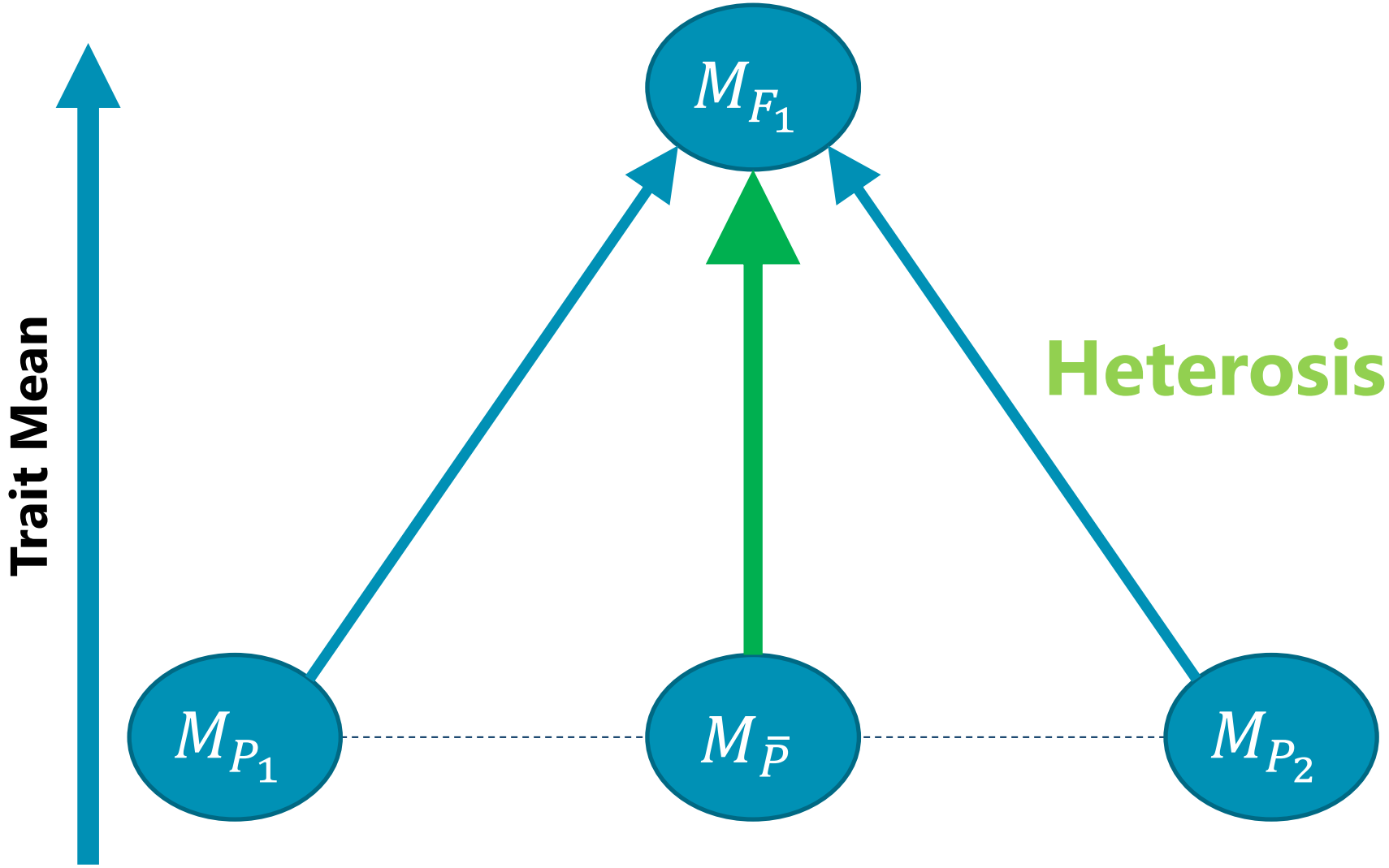
- Exploited with crossbreeding
  - Terminal crossbreeding similar to maize
- Results aren’t as drastic as in maize
  - Animals aren’t fully inbred
    - Lines/breeds are somewhat inbred
  - Exploiting panmictic midparent heterosis

Two Breeds or Lines (HWE)

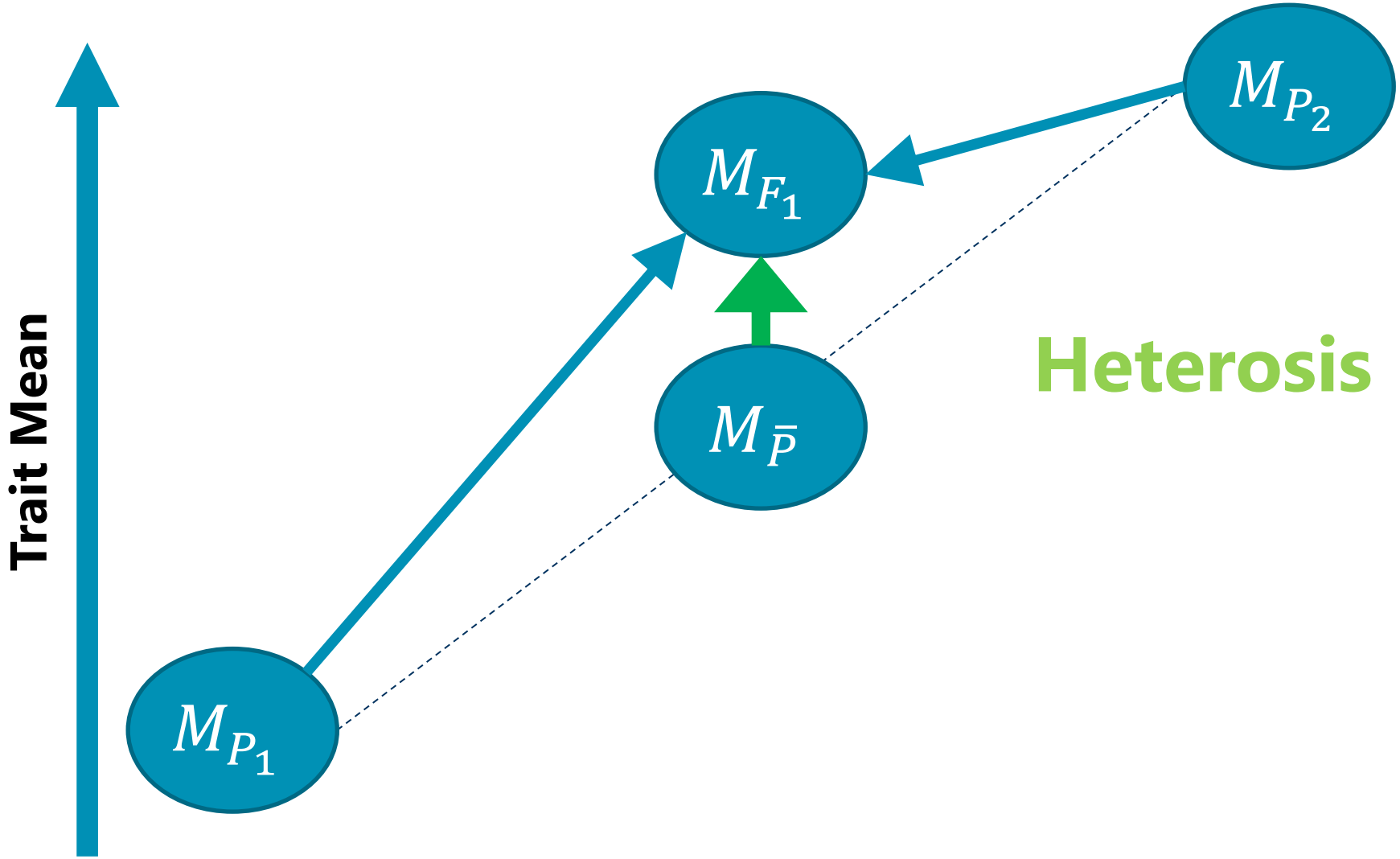
Midparent Value



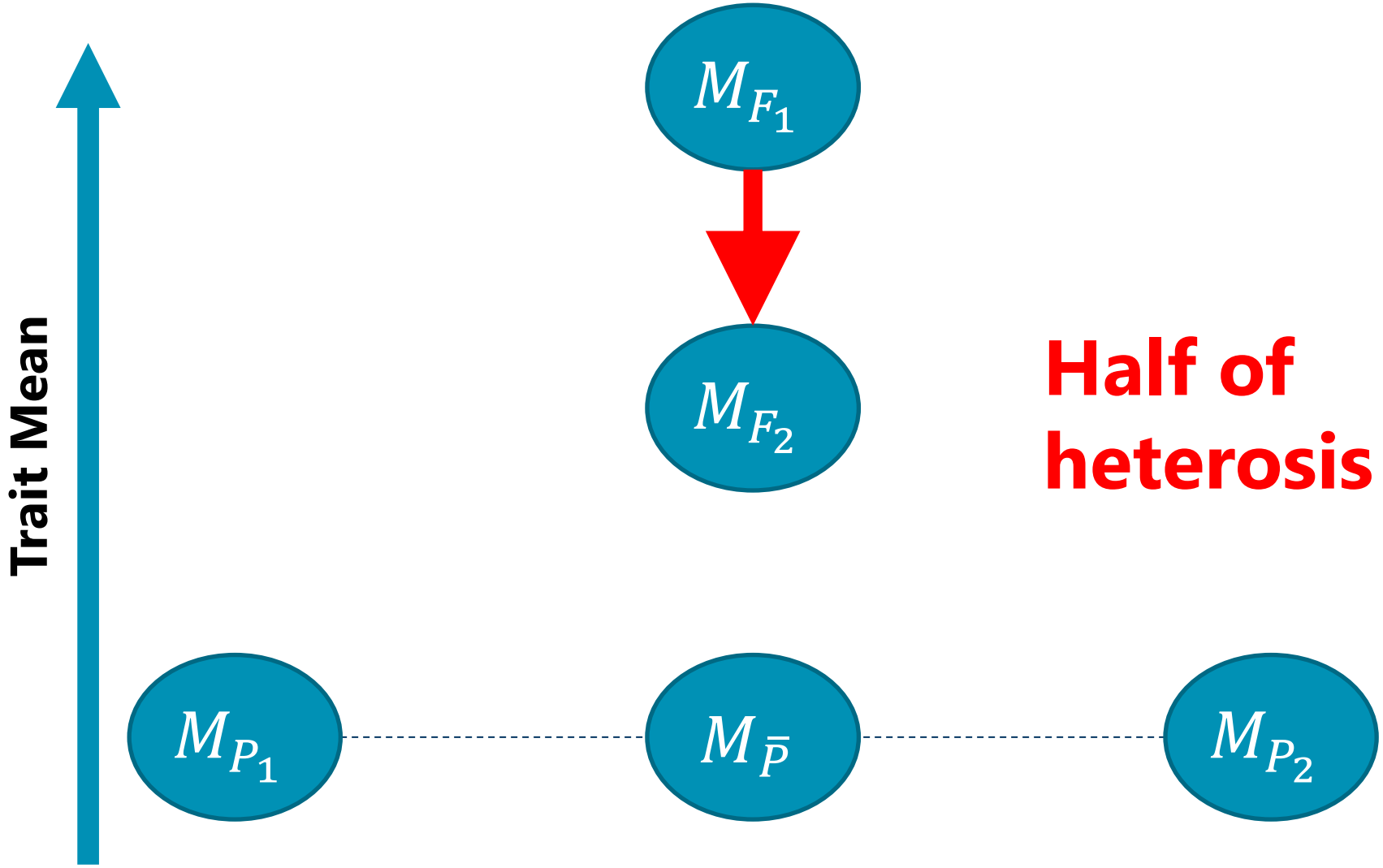
Crossbred Animals



Unequal Means



Intermate Crossbreds (HWE)



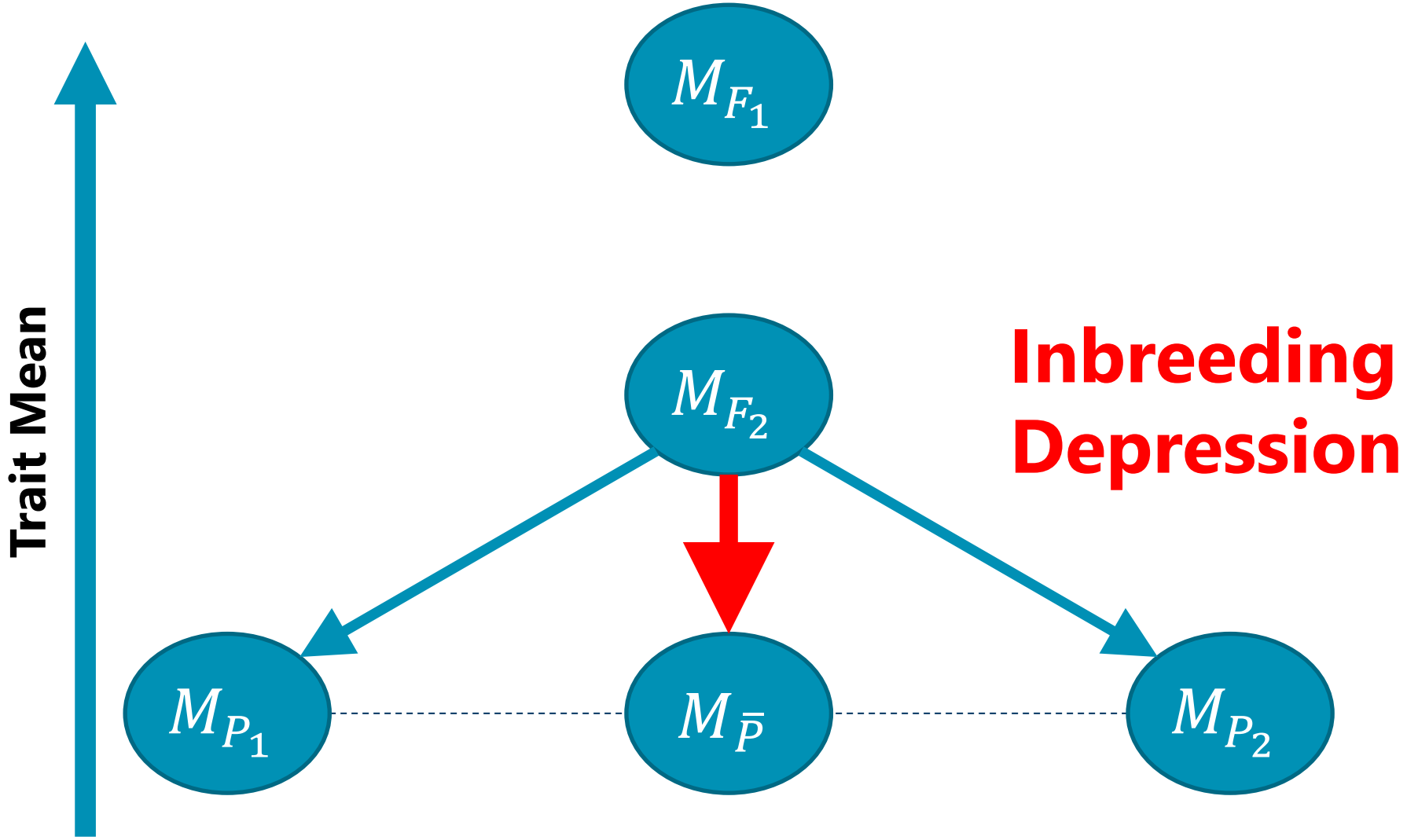
Inbreeding Depression in Animals

- Important for all breeding programs
  - Avoid mating relatives

- Textbook example of inbreeding
  - Split population into lines
  - The lines become inbred

- **Key point**: inbreeding is a relative value

Line Breeding (without Selection)



Inbreeding Depression

Population mean (no epistasis)

$$M_F = \sum (p - q)a + 2pq(1 - F)d$$

Decrease in mean due to inbreeding

$$I = M_{F=0} - M_{F=1} = \sum 2pqd$$

Heterosis

Population means

$$M_{P_1} = \sum (p - q)a + 2pqd$$

$$M_{P_2} = \sum (p' - q')a + 2p'q'd$$

Panmictic Midparent Heterosis

$$H = M_{F_1} - M_{\bar{P}} = \sum (p - p')^2 d = \sum 4\bar{p}\bar{q}Fd$$

Simulating QTL effects

Additive effects

$a \sim N(0, \sigma_a^2)$

Dominance effects

$d = \delta|a|$

Dominance degrees

$\delta \sim N(\mu_\delta, \sigma_\delta^2)$

Mean of dominance effects

$E(d) = E(\delta)E(|a|) = \mu_\delta \sigma_a \sqrt{2/\pi}$

Inbreeding Depression in Simulation

$$E(I) = \sum 2pq\boldsymbol{E}(d) = \mu_\delta \sigma_a \sqrt{2/\pi} \sum 2pq$$

- Depends on user supplied parameters
  - Mean dominance degree
  - Number of QTL
- Depends on additive effect variance

Heterosis in Simulation

$$E(H) = \sum (p - p')^2 E(d) = \mu_\delta \sigma_a \sqrt{2/\pi} \sum (p - p')^2$$

- Depends on user supplied parameters
  - Mean dominance degree
  - Number of QTL
  - Allele frequency difference

- Depends on additive effect variance

Additive Effect Variance (σ_a^2)

- Tuned for additive genetic variance (V_A)
  - User specified value

- Accomplished using linear scaling
  1. Sample effects from a standard normal
  2. Calculate variance
  3. Calculate and apply scaling constant

Behind the Scenes

Variance scaled using a single breed (HWE)

Additive genetic variance

$$V_A = 2\sum pq\alpha^2 + 4\sum_{i<j} D_{i,j}\alpha_i\alpha_j$$

Average effect

$$\alpha = a + d(q - p)$$

Average Effect Expectations

$$E(\alpha^2) = E(a^2) + E(d^2)(q-p)^2 + 2E(ad)(q-p)$$

$$= \sigma_a^2\left[1 + \left(\mu_\delta^2 + \sigma_\delta^2\right)(q-p)^2\right]$$

$$E(\alpha_1\alpha_2) = E(a_1)E(a_2) + \cdots + E(d_1)E(d_2)(q_1-p_1)(q_2-p_2)$$

$$= \sigma_a^2\mu_\delta^2\left(\frac{2}{\pi}\right)(q_1-p_1)(q_2-p_2)$$

Expectation for σ_a^2

$$E(\sigma_a^2) = \frac{V_A}{C_1 + (\mu_\delta^2 + \sigma_\delta^2)C_2 + \mu_\delta^2 C_3}$$

$$C_1 = 2\sum pq$$

$$C_2 = 2\sum pq(q-p)^2$$

$$C_3 = \frac{8}{\pi}\sum_{i<j}(q_i - p_i)(q_j - p_j)D_{i,j}$$

Expectation for Inbreeding Depression

$$E(I) = 2\sum pqE(d) = \mu_\delta \sigma_a C_1 \sqrt{2/\pi}$$

$$E(I) = \frac{\mu_\delta C_1 \sqrt{2V_A/\pi}}{\sqrt{C_1 + \left(\mu_\delta^2 + \sigma_\delta^2\right)C_2 + \mu_\delta^2 C_3}}$$

- Expectation for heterosis is similar

Deciphering the Equation

$$E(I) = \frac{\mu_\delta C_1 \sqrt{2V_A/\pi}}{\sqrt{C_1 + \left(\mu_\delta^2 + \sigma_\delta^2\right)C_2 + \mu_\delta^2 C_3}}$$

- Increasing V_A increases inbreeding depression
  - Use $\frac{I}{\sqrt{V_A}}$ for making comparisons

- Increasing σ_δ^2 decreases inbreeding depression

Simplifying Assuming p = 0.5

$$E(I) = \mu_\delta \sqrt{\frac{V_A n_{QTL}}{\pi}}$$

$$E(I) \propto \mu_\delta$$

$$E(I) \propto \sqrt{n_{QTL}}$$

Key point: trade-off between μ_δ and n_{QTL}

Practical Application

- No unique solution

- Narrow search space if possible
  - Dominance variance
  - Purebred-crossbred correlations

- Test a range of values
  - Checks sensitive to assumptions
  - Use the equations to help with this

Concluding Comments

- Incorporating dominance gets tricky

- Make sure doing so is necessary
  - Strategies for avoiding inbreeding depression
  - Strategies for exploiting heterosis

- Remember all models are wrong
  - We are seeking a useful model

AlphaSimR Demonstration