



LOUISIANA STATE UNIVERSITY
College of Agriculture
School of Plant, Environmental, and Soil Sciences
AGRO 7075 Prediction-based Breeding



Base populations and breeding schemes

Prof. Roberto Fritsche-Neto

rfneto@agcenter.lsu.edu

Baton Rouge, May 1st, 2023

Building a base population

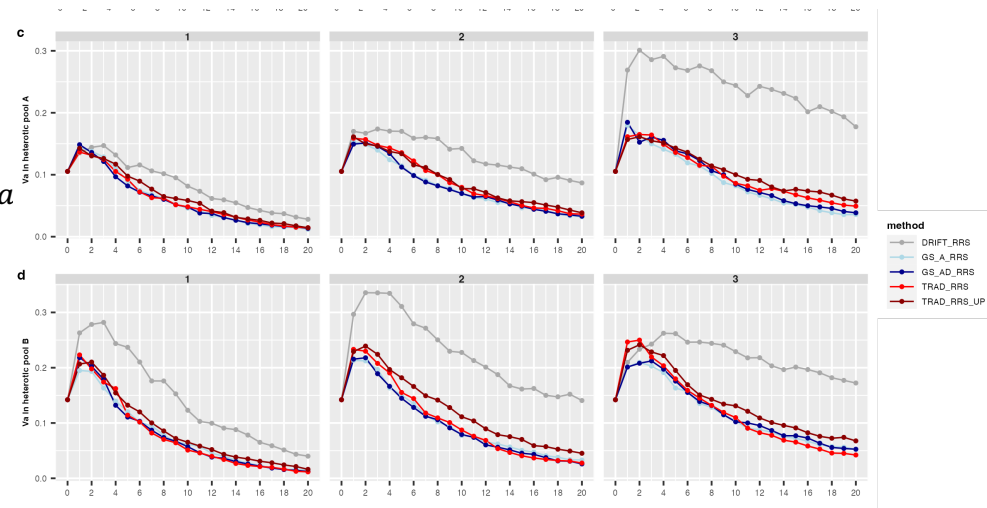
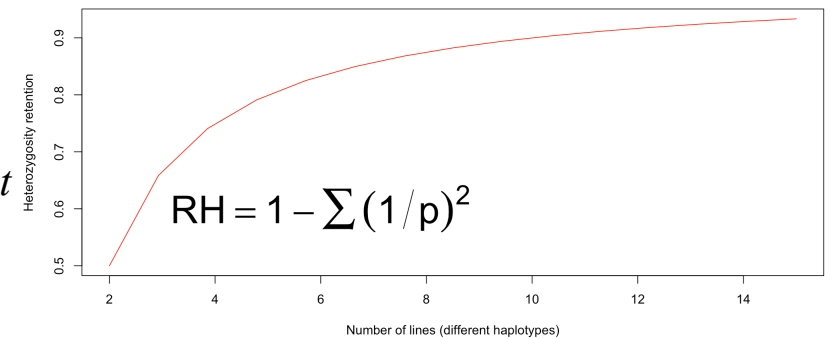
- Combine high mean and large (and good) variability
- Open mating among parents
- Closed system – *since the begin, all the alleles are present*
- HWE - binomial and multinomial distributions

- **How many parents?**

- Variability vs. Mean vs. % Heterozygosity
- Number of cycles vs. Ne
- What is the maximum of RS? $RS = 2 \cdot Ne \cdot i \cdot r_{aP} \cdot \sigma_a$
- Choosing parents:

- *Breeding program objectives*

- *Heterotic Groups (SCA and GCA)*



What is the best base population?

- Combine mean and variability (divergence)
- $L_1 = AAB BCC$
- $L_2 = aab bCC$
- $L_1 = u + \alpha_a + \alpha_b + \alpha_c$
- $L_2 = u - \alpha_a - \alpha_b + \alpha_c$
- $X_{P(12)} = u + \alpha_c$
- Which is the best populations to obtain lines?
- $X_S = u + (p - q)a + 2pqd - 2pqdF$
- $X_{S0} = u + (p - q)a + 2pqd$
- $X_{S1} = u + (p - q)a + pqd$
- The best can be estimated by $= 2.X_{S1} - X_{S0}$
- $2u + 2(p - q)a + 2pqd - (u + (p - q)a + 2pqd)$
- $u + (p - q)a$ (estimated value for $F = 1$)
- $L_1 = AAB BCC$
- $L_2 = AAb bCC$
- $L_3 = aaB BCC$
- $L_4 = aab bCC$
- $L_1 = u + \alpha_a + \alpha_b + \alpha_c$
- $L_2 = u + \alpha_a - \alpha_b + \alpha_c$
- $L_3 = u - \alpha_a + \alpha_b + \alpha_c$
- $L_4 = u - \alpha_a - \alpha_b + \alpha_c$
- $X_{P(1234)} = u + \alpha_c$

Why after so many cycles there is V_a ?

- Two-locus model = $A_i A_j / B_k B_l$
- $G_{ijkl} = (\alpha_i + \alpha_j + S_{ij}) + (\alpha_k + \alpha_l + S_{kl}) + I_{ijkl}$
- $G_{ijkl} = u + A_{ij} + B_{kl} + I_{ijkl}$
- The epistatic effect is
- $I_{ijkl} = G_{ijkl} - (u + A_{ij} + B_{kl})$
- 1 or 2 dominant allele = 13
- Otherwise = 1
- $p = q = 0.5$
- Population mean
- $u = p^2 p^2 G_{1111} + p^2 2pq G_{1112} + \dots + q^2 q^2 G_{2222}$
- $= \frac{1}{4} \cdot \frac{1}{4} \cdot 13 + \frac{1}{4} \cdot \frac{1}{2} \cdot 13 + \dots + \frac{1}{4} \cdot \frac{1}{4} \cdot 1 = 12\frac{1}{4}$
- $A_i A_j$ and $B_k B_l$ effects
- $G_{22..} = p^2 13 + 2pq 13 + q^2 1 = 10$
- $A_{22} = G_{ij..} - u = 10 - 12\frac{1}{4} = -9/4$
- I_{ijkl} effect
- $I_{1111} = G_{1111} - (u + A_{11} + B_{11})$
- $= 13 - (12 + \frac{1}{4}) - \frac{3}{4} - \frac{3}{4} = -\frac{3}{4}$

	B ₁ B ₁	B ₁ B ₂	B ₂ B ₂	G _{ij..}	A _{ij..}
A ₁ A ₁	G ₁₁₁₁ = 13	G ₁₁₁₂ = 13	G ₁₁₂₂ = 13	G _{11..} = 13	A ₁₁ = ³ / ₄
	u+A ₁₁ +B ₁₁ = 13 ³ / ₄	u+A ₁₁ +B ₁₂ = 13 ³ / ₄	u+A ₁₁ +B ₂₂ = 10 ³ / ₄		
	I ₁₁₁₁ = - ³ / ₄	I ₁₁₁₂ = - ³ / ₄	I ₁₁₂₂ = 9/4		
A ₁ A ₂	G ₁₂₁₁ = 13	G ₁₂₁₂ = 13	G ₁₂₂₂ = 13	G _{12..} = 13	A ₁₂ = ³ / ₄
	u+A ₁₂ +B ₁₁ = 13 ³ / ₄	u+A ₁₂ +B ₁₂ = 13 ³ / ₄	u+A ₁₂ +B ₂₂ = 10 ³ / ₄		
	I ₁₂₁₁ = - ³ / ₄	I ₁₂₁₂ = - ³ / ₄	I ₁₂₂₂ = 9/4		
A ₂ A ₂	G ₂₂₁₁ = 13	G ₂₂₂₁ = 13	G ₂₂₂₂ = 1	G _{22..} = 10	A ₂₂ = -9/4
	u+A ₂₂ +B ₁₁ = 10 ³ / ₄	u+A ₂₂ +B ₂₁ = 10 ³ / ₄	u+A ₂₂ +B ₂₂ = 7 ³ / ₄		
	I ₂₂₁₁ = 9/4	I ₂₂₂₁ = 9/4	I ₂₂₂₂ = -27/4		
G _{..kl}	G _{..11} = 13	G _{..12} = 13	G _{..22} = 10	u = 12 ¹ / ₄	
B _{kl}	B ₁₁ = ³ / ₄	B ₁₂ = ³ / ₄	B ₂₂ = -9/4		

Why after so many cycles there is V_a ?

- **Considering** $AA = Aa = 13$ and $aa = 10$
- $a = d = 3/2$
- $V_a = 2pq[a + (p - q)d]^2$
 - $= 2 \cdot \frac{1}{2} \cdot \frac{1}{2} [3/2 + (\frac{1}{2} - \frac{1}{2})3/2]$
 - $= 9/8 = 1.12$
- $V_d = (2pqd)^2$
 - $= (2 \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 3/2)^2$
 - $= 9/16 = 0.56$
- **The same values are found for B. Thus,**
- $V_a = V_{aA} + V_{aB} = 2.25$
- $V_d = V_{dA} = V_{dB} = 1.12$
- $V_g = \frac{1}{4} \frac{1}{4} (13 - (12 + 1/4))^2 + \dots + \frac{1}{4} \frac{1}{4} (1 - 12 \frac{1}{4})^2$
 - $= 135/16 = 8.44$
- $V_i = V_g - V_a - V_d = 81/16 = 5.06$
- $V_a = 13.33 \%$
- $V_d = 6.66 \%$
- $V_i = 80.01 \%$

	B ₁ B ₁	B ₁ B ₂	B ₂ B ₂	G _{ij..}	A _{ij..}
A ₁ A ₁	G ₁₁₁₁ = 13	G ₁₁₁₂ = 13	G ₁₁₂₂ = 13	G _{11..} = 13	A ₁₁ = ¾
	u+A ₁₁ +B ₁₁ = 13¾	u+A ₁₁ +B ₁₂ = 13¾	u+A ₁₁ +B ₂₂ = 10¾		
	I ₁₁₁₁ = -¾	I ₁₁₁₂ = -¾	I ₁₁₂₂ = 9/4		
A ₁ A ₂	G ₁₂₁₁ = 13	G ₁₂₁₂ = 13	G ₁₂₂₂ = 13	G _{12..} = 13	A ₁₂ = ¾
	u+A ₁₂ +B ₁₁ = 13¾	u+A ₁₂ +B ₁₂ = 13¾	u+A ₁₂ +B ₂₂ = 10¾		
	I ₁₂₁₁ = -¾	I ₁₂₁₂ = -¾	I ₁₂₂₂ = 9/4		
A ₂ A ₂	G ₂₂₁₁ = 13	G ₂₂₂₁ = 13	G ₂₂₂₂ = 1	G _{22..} = 10	A ₂₂ = -9/4
	u+A ₂₂ +B ₁₁ = 10¾	u+A ₂₂ +B ₂₁ = 10¾	u+A ₂₂ +B ₂₂ = 7¾		
	I ₂₂₁₁ = 9/4	I ₂₂₂₁ = 9/4	I ₂₂₂₂ = -27/4		
G _{..kl}	G _{..11} = 13	G _{..12} = 13	G _{..22} = 10	u = 12¾	
B _{kl}	B ₁₁ = ¾	B ₁₂ = ¾	B ₂₂ = -9/4		

Why after so many cycles there is Va?

- Considering B_2B_2 fixed, $AA = Aa = 13$, and $aa = 10$
- $u = p^2G_{1122} + 2pqG_{1222} + q^2G_{2222}$
- $= \frac{1}{4}.13 + 2.\frac{1}{2}.\frac{1}{2}.13 + \frac{1}{4}.1 = 10$
- A_iA_j effect
- $G_{22..} = 1$
- $A_{22} = G_{ij..} - u = 1 - 10 = -9$
- $I_{ijkl} = G_{ijkl} - (u + A_{ij} + B_{kl})$
- $I_{1111} = 13 - 10 - 3 + 0 = 0$
- **Variances**
- $Vg = \frac{1}{4}(13 - 10)^2 + 2.\frac{1}{2}.\frac{1}{2}(13 - 10)^2 + \frac{1}{4}(1 - 10)^2 = 27$
- $a = d = 6$
- $Va = 2pq[a + (p - q)d]^2 = 18 = 66.66\%$
- $Vd = (2pqd)^2 = 9 = 33.33\%$
- $Vi = Vg - Va - Vd = 0 = 0\%$

	B_2B_2	$G_{ij..}$	$A_{ij..}$
A_1A_1	$G_{1122} = 13$	$G_{11..} = 13$	$A_{11} = 3$
	$u+A_{11}+B_{22}$		
	$I_{1122} = 0$		
A_1A_2	$G_{1222} = 13$	$G_{12..} = 13$	$A_{12} = 3$
	$u+A_{12}+B_{22}$		
	$I_{1222} = 0$		
A_2A_2	$G_{2222} = 1$	$G_{22..} = 1$	$A_{22} = -9$
	$u+A_{22}+B_{22}$		
	$I_{2222} = 0$		
		$u = 10$	

Main criteria to choose the breeding method

- **Propagation system** - *Sexual (cross-pollination or self-pollination) or clonal propagated*
- **Trait** - *Qualitative vs. Quantitative*
- **Heritability** - *Low or high*
- **Genetic control** - *Additive vs. non-additive*
- **Proportion explored of the additive genetic variance** – *Cov between evaluated and improved population*
- **Resources available** – *time, money, labor, ...*
- **Breeding stage**: *F2, F3, etc.*
- **Product to be developed** - *lines, hybrid, variety, etc.*

How can we maximize the genetic gain per time?

- Increasing the intensity of selection (i)
- Minimizing the environment effect – **replicates and randomize**

$$\sigma_P^2 = \sigma_G^2 + \sigma_E^2 + 2COV_{GE} \quad h_A^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2/r}$$

$$RS = i \cdot r_{aP} \cdot \sigma_a / T$$

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{\sigma_a^2}{\sigma_a \sigma_P} = \frac{\sigma_a}{\sigma_P}$$

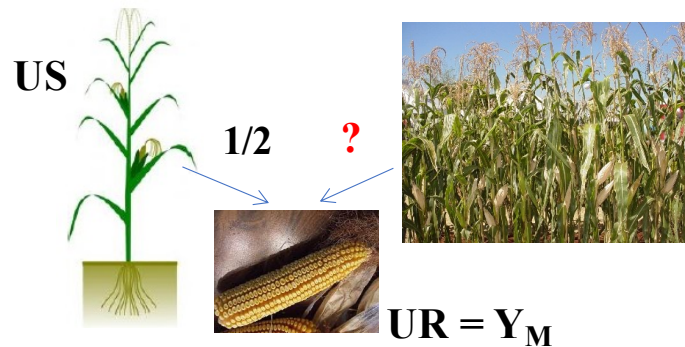
- Increasing the elite genetic variability
- **Using more accurate or less time-consuming breeding schemes**
- *For instance, combining GS and HTP*
- Maximizing the additive genetic covariance between evaluated and improved population
- **c** = pathway between the unit of selection (US) and improved population (Y_M) – **indirect selection**

$$RS = i \frac{\sigma_a}{\sigma_P} \cdot \sigma_a \quad RS = i \frac{\sigma_a^2}{\sigma_P}$$

$$RS(us, Y_m) = i \frac{\sigma_{GY_m}^2}{\sigma_{PY_m}} \cdot \frac{COV_{G(us, Y_m)}}{\sigma_{GY_m}^2} \quad RS = i \frac{COV_{G(us, Y_m)}}{\sigma_{PY_m}} \quad RS = i \frac{r_{g(us, Y_m)} \sigma_{a(us)} \sigma_{a(Y_m)}}{\sigma_{PY_m}} \quad RS = i \cdot c \cdot \frac{\sigma_a^2}{\sigma_{PY_m}}$$

Massal selection

- Harvest together similar and superior phenotypes
- They will form the newest improved population
- Only one sex (female)



- Little gains
- Easy, cheap, and no time-consuming
- Species little improved and high heritable traits

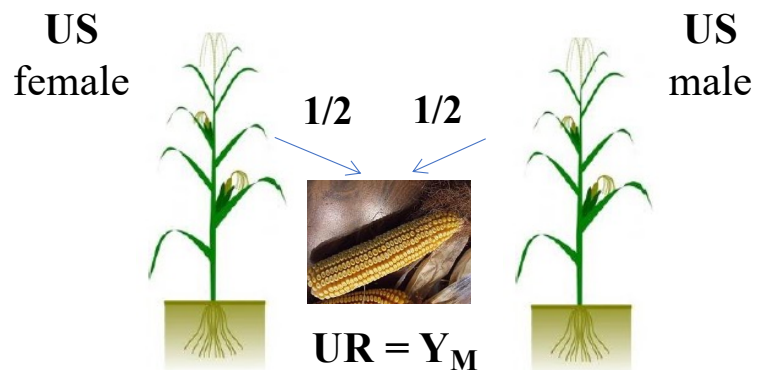
$$RS = i. c. \frac{\sigma_a^2}{\sigma_{PYm}}$$

$$RS = i. \frac{1}{2} \cdot \frac{\sigma_a^2}{\sigma_{PYm}}$$

$$RS = i. \frac{1}{2} \cdot \frac{\sigma_a^2}{\sqrt{\sigma_G^2 + \sigma_E^2 + 2COV_{GE}}}$$

Massal selection

- Both parents



$$RS = i_1 \cdot \frac{1}{2} \cdot \frac{\sigma_a^2}{\sigma_{PYm}} + i_2 \cdot \frac{1}{2} \cdot \frac{\sigma_a^2}{\sigma_{PYm}}$$

$$RS = i \cdot \frac{\sigma_a^2}{\sigma_{PYm}}$$

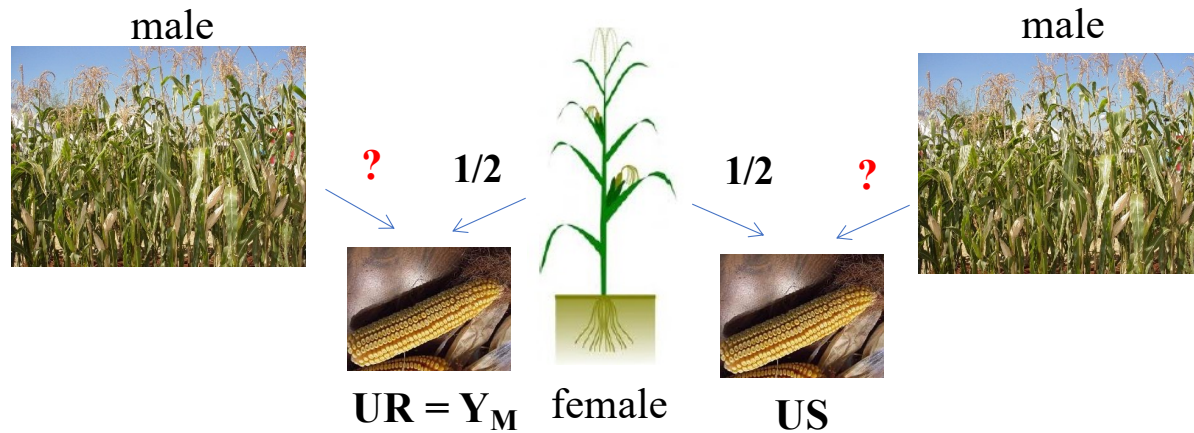
$$RS = i_1 \cdot c \cdot \frac{\sigma_a^2}{\sigma_{PYm}} + i_2 \cdot c \cdot \frac{\sigma_a^2}{\sigma_{PYm}}$$

$$RS = i \cdot \frac{\sigma_a^2}{\sqrt{\sigma_G^2 + \sigma_E^2 + 2COV_{GE}}}$$

- Better gains but still low

Selection based on progenies

- Among half-sibs (**only one sex**)



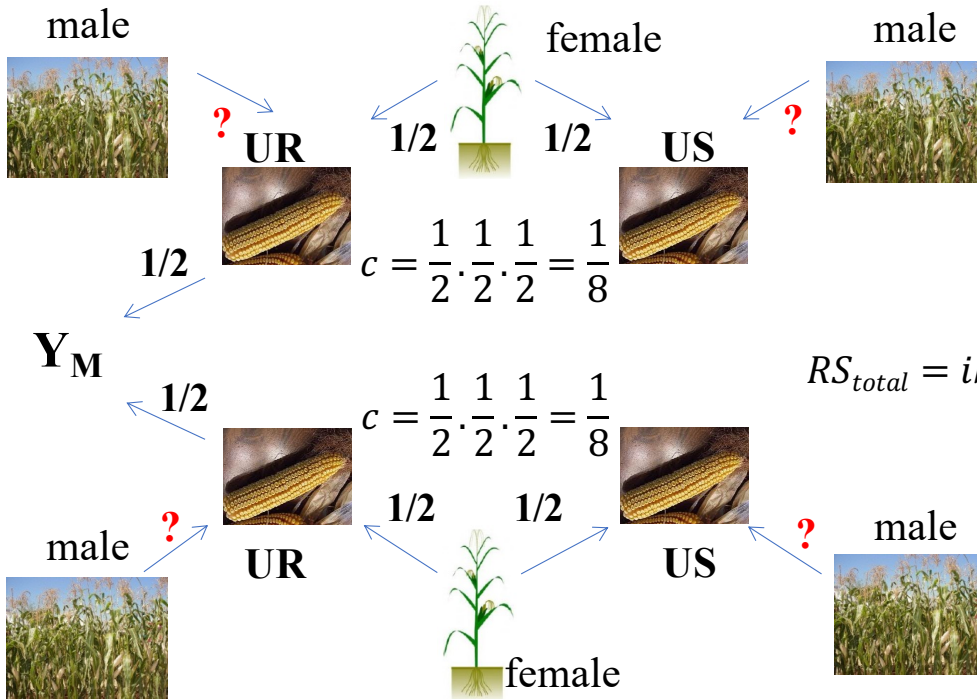
$$c = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$RS = i \cdot \frac{1}{4} \cdot \frac{\sigma_a^2}{\sigma_{PYm}}$$

$$RS = i \cdot \frac{1}{4} \cdot \frac{\sigma_a^2}{\sqrt{\sigma_G^2 + \sigma_E^2}/r}$$

Selection based on progenies

- Half-sibs but in both parents

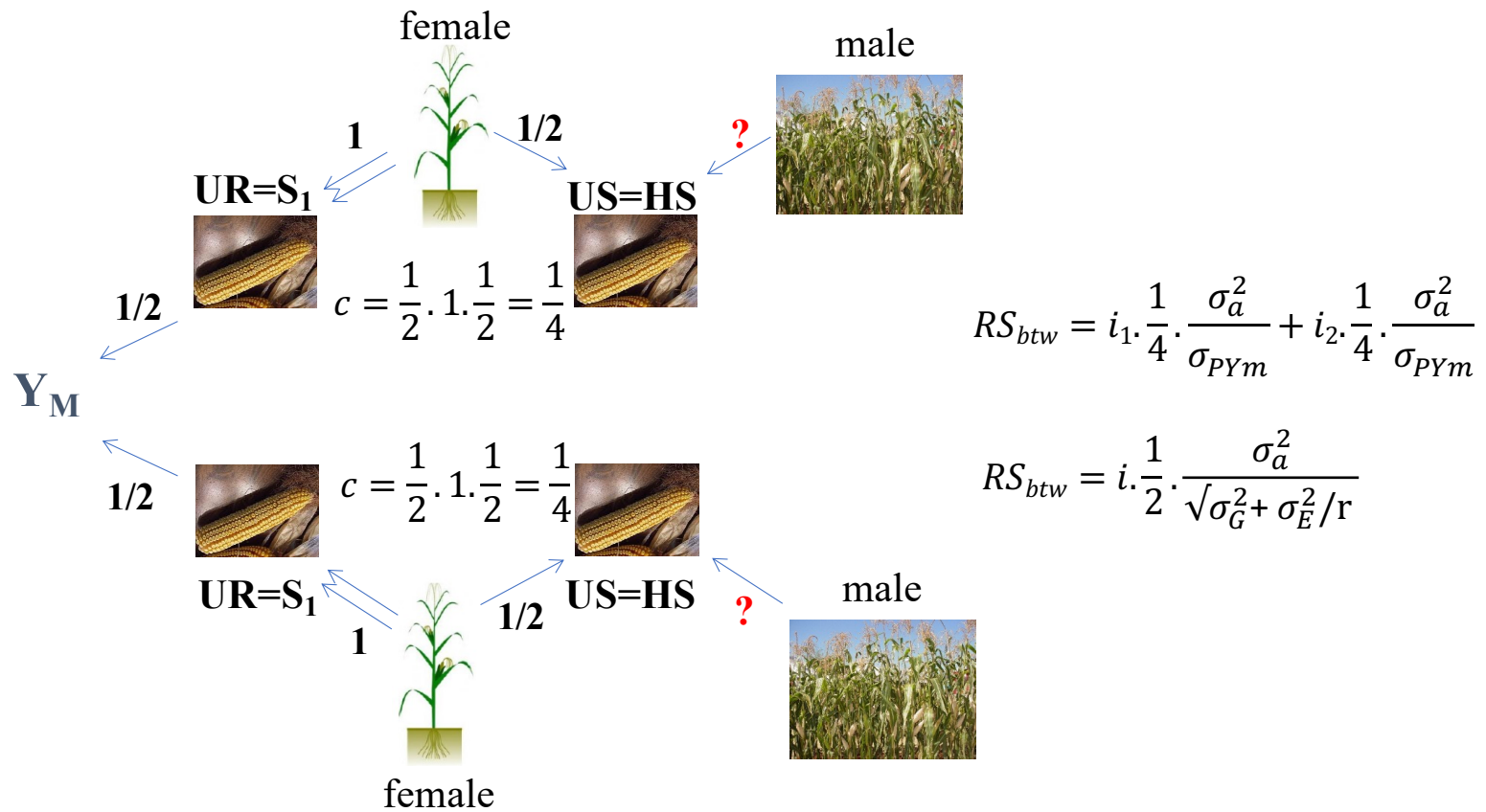


$$RS_{btw} = i_1 \cdot \frac{1}{8} \cdot \frac{\sigma_a^2}{\sigma_{PYm}} + i_2 \cdot \frac{1}{8} \cdot \frac{\sigma_a^2}{\sigma_{PYm}}$$

$$RS_{btw} = i \cdot \frac{1}{4} \cdot \frac{\sigma_a^2}{\sigma_{PYm}} \quad RS_{btw} = i \cdot \frac{1}{4} \cdot \frac{\sigma_a^2}{\sqrt{\sigma_G^2 + \sigma_E^2/r}}$$

$$RS_{total} = ibtw. \frac{1}{4} \cdot \frac{\sigma_a^2}{\sqrt{\sigma_G^2 + \sigma_E^2/r}} + i_{wth.} \frac{3}{4} \cdot \frac{\sigma_a^2}{\sqrt{\frac{3}{4}\sigma_G^2 + \sigma_D^2 + \sigma_E^2 + \sigma_d^2}}$$

Recurrent selection – HS / S₁



Comparison between breeding methods

- Response to selection per time

$$RS = \frac{i \cdot r_{aP} \cdot \sigma_a}{T} \qquad EF_{X/Y} = \frac{RS_X \cdot T_Y}{RS_Y \cdot T_X} \cdot 100$$

- It is possible simulate innumerous breeding schemes
- Moreover, the cost per unit / plot must be considered
- Thus, there is no “the best method”