UNIVERSIDADE DE SÃO PAULO
ESCOLA SUPERIOR DE AGRICULTURA
"LUIZ DE QUEIROZ"
DEPARTAMENTO DE GENÉTICA
LGN5825 Genética e Melhoramento de Espécies Alógamas

Base populations and breeding schemes

Prof. Roberto Fritsche-Neto

roberto.neto@usp.br

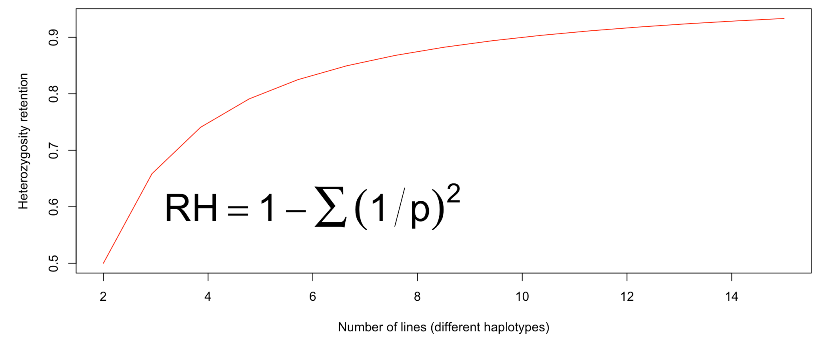
Piracicaba, October 23rd, 2019

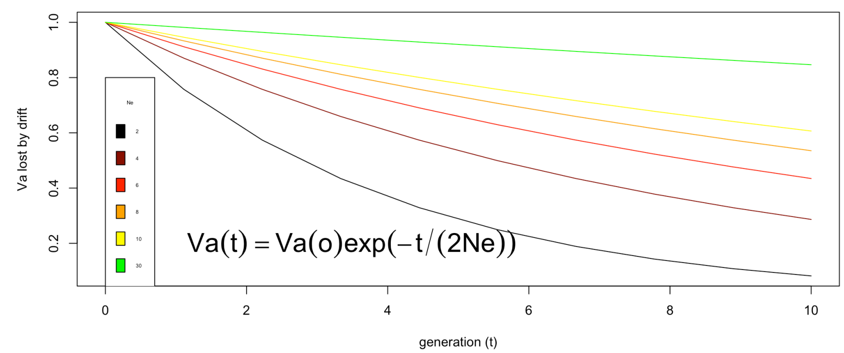
Building a base population

- Combine high mean and large 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.* % Heterosigozyty
- Number of cycles *vs.* Ne
- What is the maximum of RS? $RS = 2.Ne.i.r_{aP}.\sigma_a$
- Choosing parents:

 - *Breeding program objectives*

 - *Heterotic Groups (SCA and GCA)*





What is the best base population?

- Combine mean and variability (divergence)
- $L_1 = AABBCC$
- $L_2 = aabbCC$
- $L_1 = u + \boldsymbol{\alpha}_a + \boldsymbol{\alpha}_b + \boldsymbol{\alpha}_c$
- $L_2 = u - \boldsymbol{\alpha}_a - \boldsymbol{\alpha}_b + \boldsymbol{\alpha}_c$
- $X_{P(12)} = u + \boldsymbol{\alpha}_c$

- Which are 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 $F = 1$)

- $L_1 = AABBCC$
- $L_2 = AAbbCC$
- $L_3 = aaBBCC$
- $L_4 = aabbCC$

- $L_1 = u + \boldsymbol{\alpha}_a + \boldsymbol{\alpha}_b + \boldsymbol{\alpha}_c$
- $L_2 = u + \boldsymbol{\alpha}_a - \boldsymbol{\alpha}_b + \boldsymbol{\alpha}_c$
- $L_3 = u - \boldsymbol{\alpha}_a + \boldsymbol{\alpha}_b + \boldsymbol{\alpha}_c$
- $L_4 = u - \boldsymbol{\alpha}_a - \boldsymbol{\alpha}_b + \boldsymbol{\alpha}_c$
- $X_{P(1234)} = u + \boldsymbol{\alpha}_c$

Why after so many cycles there is Va?

- Two-locus model = A_iA_j / B_kB_l
- $G_{ijkl} = (\boldsymbol{\alpha}_i + \boldsymbol{\alpha}_j + \boldsymbol{S}_{ij}) + (\boldsymbol{\alpha}_k + \boldsymbol{\alpha}_l + \boldsymbol{S}_{kl}) + I_{ijkl}$
- $G_{ijkl} = u + A_{ij} + B_{kl} + I_{ijkl}$
- 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
- $u = p^2p^2G_{1111} + p^22pqG_{1112} + \ldots + q^2q^2G_{2222}$
- $= ¼.¼.13 + ¼.½ .13 + \ldots + ¼.¼.1 = 12¼$
- A_iA_j and B_kB_l effects
- $G_{22..} = p^2 13 + 2pq13 + q^2 1 = 10$
- $A_{22} = Gij.. - u = 10 - 12 ¼ = -9/4$
- I_{ijkl} effect
- $I_{1111} = G_{1111} - (u + A_{11} + B_{11})$
- $= 13 - (12+ ¼) - ¾ - ¾ = -¾$

	B_1B_1	B_1B_2	B_2B_2	$G_{ij..}$	$A_{ij..}$
	$G_{1111} = 13$	$G_{1112} = 13$	$G_{1122} = 13$		
A_1A_1	$u+A_{11}+B_{11} = 13¾$	$u+A_{11}+B_{12} = 13¾$	$u+A_{11}+B_{22} = 10¾$	$G_{11..} = 13$	$A_{11} = ¾$
	$I_{1111} = -¾$	$I_{1112} = -¾$	$I_{1122} = 9/4$		
	$G_{1211} = 13$	$G_{1212} = 13$	$G_{1222} = 13$		
A_1A_2	$u+A_{12}+B_{11} = 13¾$	$u+A_{12}+B_{12} = 13¾$	$u+A_{12}+B_{22} = 10¾$	$G_{12..} = 13$	$A_{12} = ¾$
	$I_{1211} = -¾$	$I_{1212} = -¾$	$I_{1222} = 9/4$		
	$G_{2211} = 13$	$G_{2221} = 13$	$G_{2222} = 1$		
A_2A_2	$u+A_{22}+B_{11} = 10¾$	$u+A_{22}+B_{21} = 10¾$	$u+A_{22}+B_{22} = 7¾$	$G_{22..} = 10$	$A_{22} = -9/4$
	$I_{2211} = 9/4$	$I_{2221} = 9/4$	$I_{2222} = -27/4$		
$G_{..kl}$	$G_{..11} = 13$	$G_{..12} = 13$	$G_{..22} = 10$	**u = 12¼**	
B_{kl}	$B_{11} = ¾$	$B_{12} = ¾$	$B_{22} = -9/4$		

Why after so many cycles there is Va?

- Considering AA = Aa = 13 and aa = 10
- $a = d = 3/2$
- $Va = 2pq[a + (p - q)d)]^2$
- $= 2\ ½\ ½[3/2 + (½ - ½)3/2)$
- $= 9/8 = 1.12$
- $Vd = (2pqd)^2$
- $= (2\ ½\ ½3/2)^2$
- $= 9/16 = 0.56$
- The same values are found for B. Thus,
- $Va = VaA + VaB = 2.25$
- $Vd = VdA = VdB = 1.12$
- $Vg = ¼¼(13 - (12 + 1/4))^2 + \ldots + ¼¼(1 - 12¼)^2$
- $= 135/16 = 8.44$
- $Vi = Vg - Va - Vd = 81/16 = 5.06$
- Va = 13.33 %
- Vd = 6.66 %
- Vi = 80.01%

	B_1B_1	B_1B_2	B_2B_2	$G_{ij..}$	$A_{ij..}$
A_1A_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_1A_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_2A_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 = 12¼**	
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
- $u = p^2G_{1122} + 2pqG_{1222} + q^2G_{2222}$
- $= ¼.13 + 2.½.½.13 + ¼.1 = 10$
- 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
- $Vg = ¼\,(13 – 10)^2 + 2.½.½(13 – 10)^2 + ¼\,(1 - 10)^2 = 27$
- $a = d = 6$
- $Va = 2pq[a + (p – q)d)]^2 = 18 = 66.66\%$
- $Vd = (2pqd)^2 = 9 = 33.33\%$
- $Vi = Vg – Va – Vd = 0 = 0\%$

	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, …*
- Product to be developed - *lines, hybrid, variety, etc.*

How can we maximize the genetic gain per time?

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

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

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

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

- Increasing the genetic variability

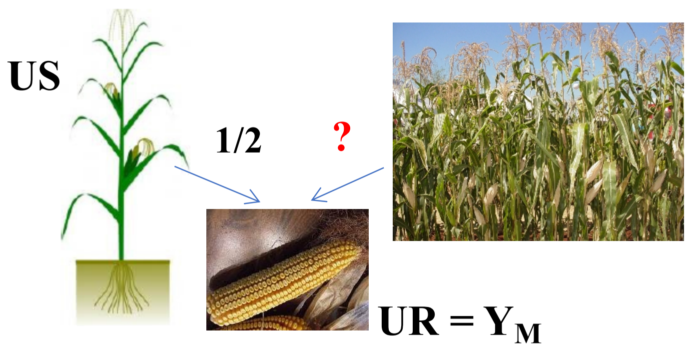
$$RS = i\frac{\sigma_a}{\sigma_P}.\sigma_a \qquad RS = i\frac{\sigma_a^2}{\sigma_P}$$

- 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

$$RS(us, Ym) = i\frac{\sigma_{GYm}^2}{\sigma_{PYm}}.\frac{COV_{G(us,Ym)}}{\sigma_{GYm}^2} \qquad RS = i\frac{COV_{G(us,Ym)}}{\sigma_{PYm}} \qquad RS = i\frac{r_{g(us,Ym)}\ \sigma_{a(us)}\sigma_{a(Ym)}}{\sigma_{PYm}} \qquad RS = i.c.\frac{\sigma_a^2}{\sigma_{PYm}}$$

Massal selection

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



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

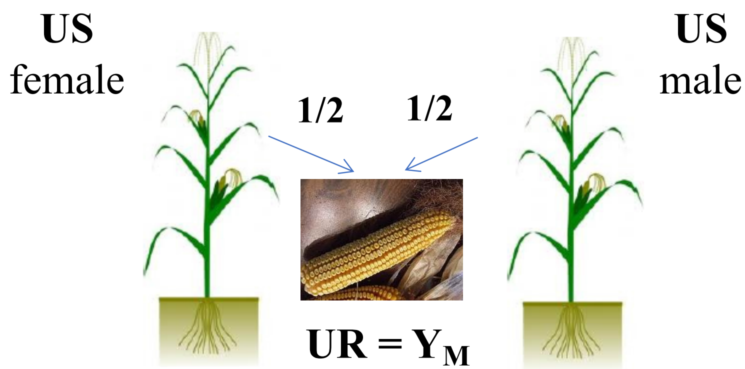
$$RS = i.\frac{1}{2}.\frac{\sigma_a^2}{\sigma_{PYm}}$$

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

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

Massal selection

- Both parents



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

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

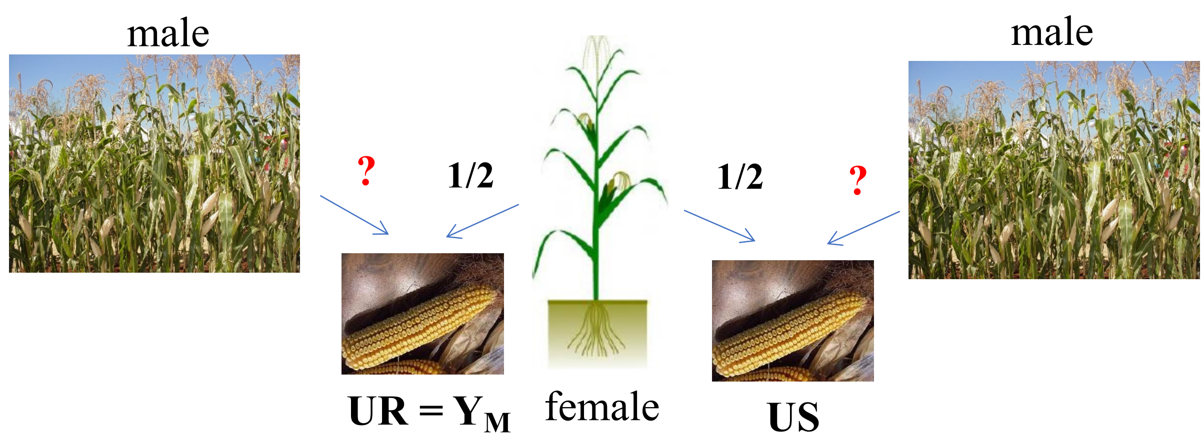
$$RS = i_1.c.\frac{\sigma_a^2}{\sigma_{PYm}} + i_2.c.\frac{\sigma_a^2}{\sigma_{PYm}}$$

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

- Better gains but still low

Selection based on progenies

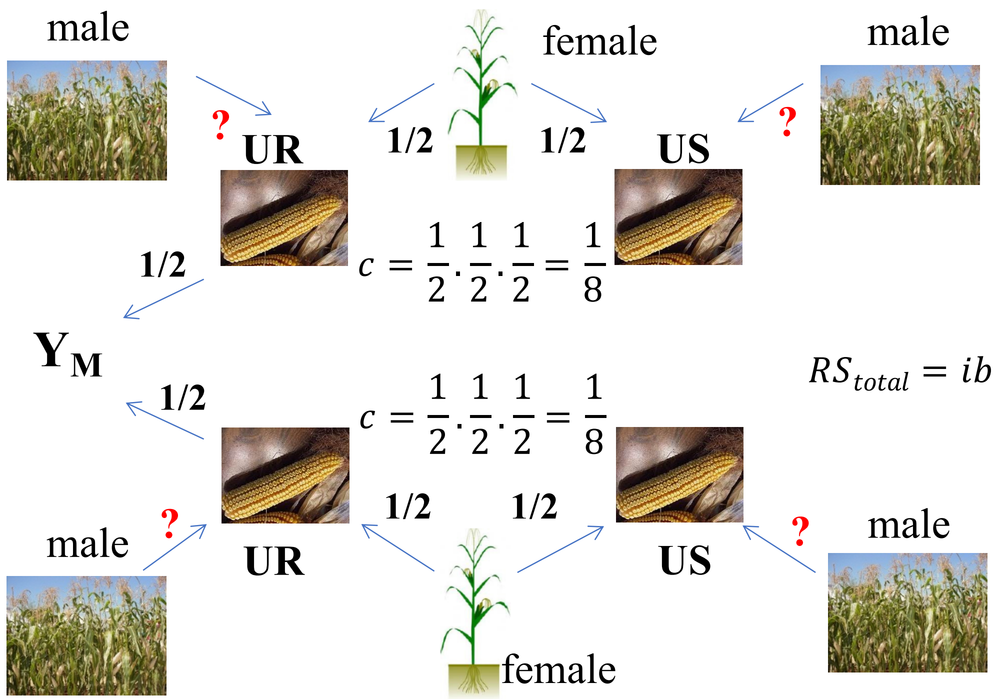
- Among half-sibs (only one sex)



$$c = \frac{1}{2} . \frac{1}{2} = \frac{1}{4} \qquad RS = i . \frac{1}{4} . \frac{\sigma_a^2}{\sigma_{PYm}} \qquad RS = i . \frac{1}{4} . \frac{\sigma_a^2}{\sqrt{\sigma_G^2 + \sigma_E^2 / \mathrm{r}}}$$

Selection based on progenies

- Half-sibs but in both parents

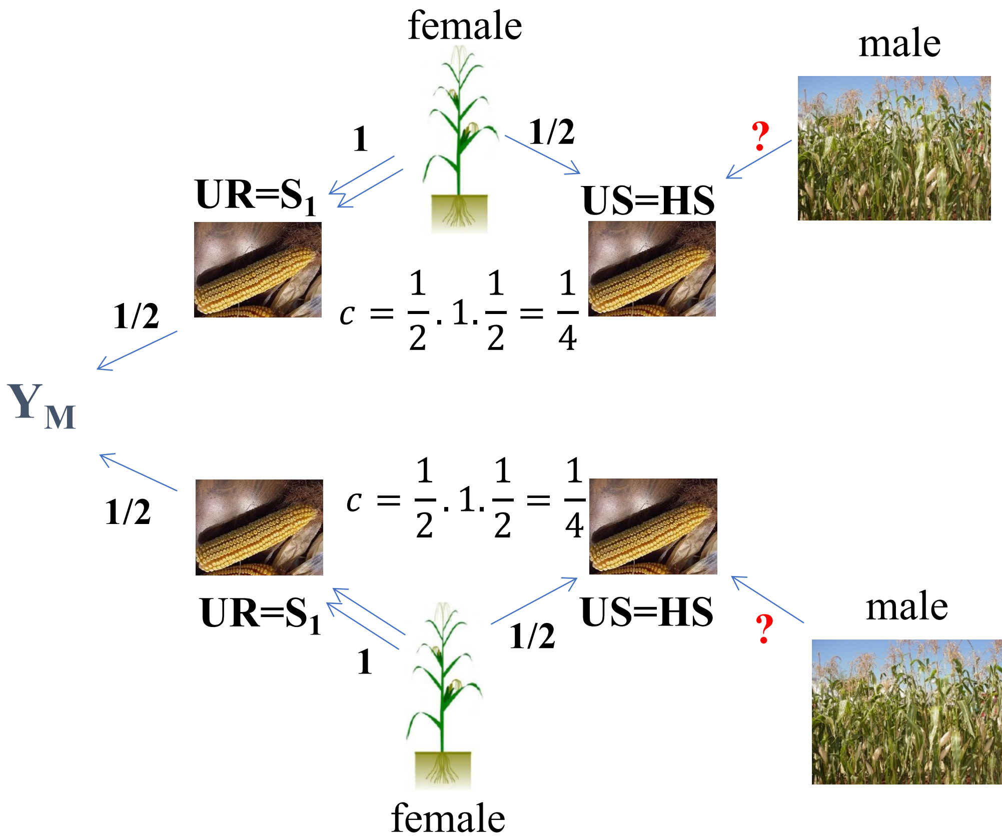


$$RS_{btw} = i_1.\frac{1}{8}.\frac{\sigma_a^2}{\sigma_{PYm}} + i_2.\frac{1}{8}.\frac{\sigma_a^2}{\sigma_{PYm}}$$

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

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

Recurrent selection – HS / S_1



$$RS_{btw} = i_1.\frac{1}{4}.\frac{\sigma_a^2}{\sigma_{PYm}} + i_2.\frac{1}{4}.\frac{\sigma_a^2}{\sigma_{PYm}}$$

$$RS_{btw} = i.\frac{1}{2}.\frac{\sigma_a^2}{\sqrt{\sigma_G^2 + \sigma_E^2/\mathrm{r}}}$$

Comparison between breeding methods

- Response to selection per time

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

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