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

Lines, testers and testcrosses

Prof. Roberto Fritsche-Neto

roberto.neto@usp.br

Piracicaba, October 16th, 2019

Main method to obtain lines

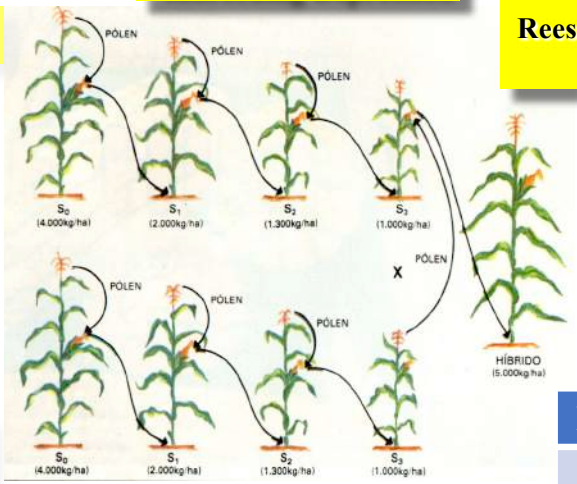
Inbreeding Depression

Reestablish the vigor Heterosis

GROUP (A)

Pedigree within group

GROUP (B)



Single-cross (12,000 Kg / ha)

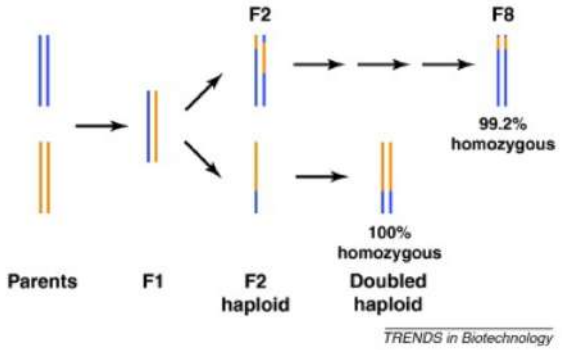
Line	Among					Within				
	Va	Vd	D1	D2	H	Va	Vd	D1	D2	H
S1	1	0.25	1	0.12	0	0.5	0.25	1	0.38	0.25
S2	1.5	0.13	2.5	0.56	0.06	0.25	0.13	0.5	0.19	0.13
S3	1.75	0.06	3.25	0.78	0.05	0.13	0.06	0.25	0.09	0.06
S6	1.97	0.01	3.95	0.97	0.01	0.02	0.02	0.03	0.01	0.01
S_∞	2	0	4	1	0	0	0	0	0	0

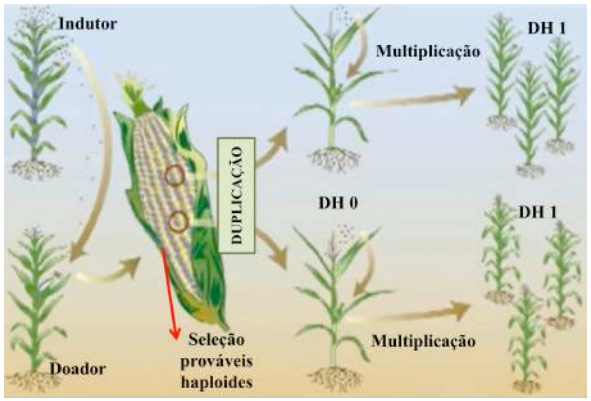
Double-haploids

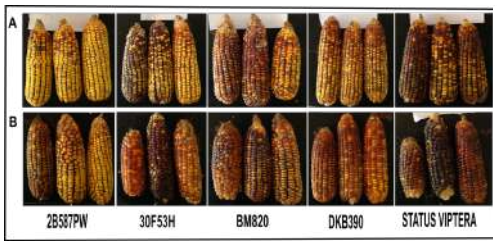
- **Advantages**
- Reduce the time to obtain lines
- The unique method to achieve F = 1
- Conserve most of the parent's haplotypes

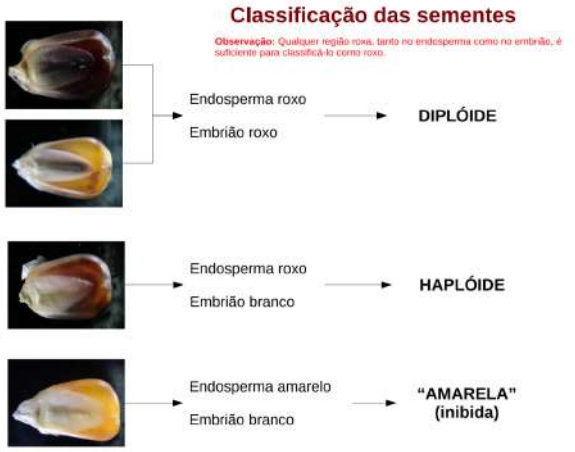
- **Drawbacks**
- It allows just one crossing-over
- There is no selection – too much variability
- Lots of lines in the end
- **Must be associated with Genomic selection**

- **Challenges**
- Low induce rate
- The identification is time-consuming and subjective
- High costs to obtain the lines
- Patents









Should we select genotypes based on lines or hybrids?

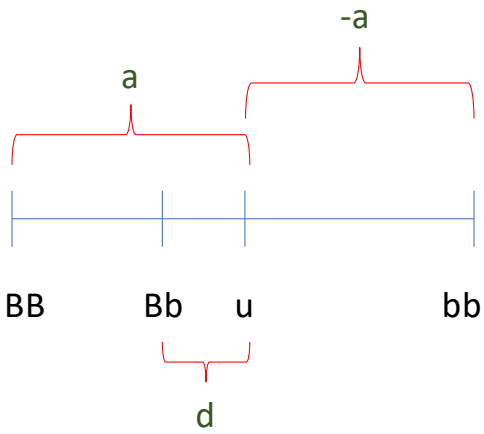
- Average degree of dominance (add)
- add = d/a = $\boldsymbol{\psi}$

- Considering an F_2 population
- p = q = 0.5
- $\boldsymbol{\alpha}$ = [a + (q – p)d]

- Va = 2pq$\boldsymbol{\alpha}^2$ = 2pqa^2 = ½a^2 $\quad a = \sqrt{2.Va}$

- Vd = $(2pqd)^2$ = ¼d^2 $\quad d = \sqrt{4.Vd}$

$$add = \frac{d}{a} = \frac{\sqrt{4.Vd}}{\sqrt{2.Va}}$$



0	Absence of dominance
0 < d/a < 1	Partial dominance
1	Complete dominance
>1	Overdominance

Correlation between lines and hybrids

$$r_{L,H} = \frac{\sigma_{LH}}{\sigma_L \ \sigma_H}$$

- Line
- $G_{ii} = \boldsymbol{\alpha}_i + \boldsymbol{\alpha}_i + \boldsymbol{S}_{ii}$
- Hybrid
- $G_{ij} = \boldsymbol{\alpha}_i + \boldsymbol{\alpha}_j + \boldsymbol{S}_{ij}$

- Genetic variance among lines
- $VgL = E[G_{ii} - E(G_{ii})]^2$
- $= E[u + \boldsymbol{\alpha}_i + \boldsymbol{\alpha}_i + \boldsymbol{S}_{ii} - u]^2$
- $= E[2\boldsymbol{\alpha}_i + \boldsymbol{S}_{ii}]^2$
- $= E[2\boldsymbol{\alpha}_i]^2 + 2E[\boldsymbol{\alpha}_i \boldsymbol{S}_{ii}] + E[\boldsymbol{S}_{ii}]^2$
- $= 4E[\boldsymbol{\alpha}_i]^2 + 2E[2\boldsymbol{\alpha}_i \boldsymbol{S}_{ii}] + E[\boldsymbol{S}_{ii}]^2$
- $= 4E[\boldsymbol{\alpha}_i]^2 + 4E[\boldsymbol{\alpha}_i \boldsymbol{S}_{ii}] + E[\boldsymbol{S}_{ii}]^2$
- $= 2Va + 4D1 + D2$

- Genetic variance among single-crosses
- $VgH = E[G_{ij} - E(G_{ij})]^2$
- $= E[u + \boldsymbol{\alpha}_i + \boldsymbol{\alpha}_j + \boldsymbol{S}_{ij} - u]^2$
- $= E[\boldsymbol{\alpha}_i + \boldsymbol{\alpha}_j + \boldsymbol{S}_{ij}]^2$
- $= E[\boldsymbol{\alpha}_i]^2 + E[\boldsymbol{\alpha}_j]^2 + E[\boldsymbol{S}_{ij}]^2 + \ldots$
- $= E[\boldsymbol{\alpha}_i]^2 + E[\boldsymbol{\alpha}_j]^2 + E[\boldsymbol{S}_{ij}]^2$
- $= ½Va + ½Va + Vd$
- $= Va + Vd$

- Covariance between lines and single-crosses
- $COV_{(L,H)} = E[G_{ij} - E(G_{ij})] \,.\, E[G_{ii} - E(G_{ii})]$
- $= E[\boldsymbol{\alpha}_i + \boldsymbol{\alpha}_j + \boldsymbol{S}_{ij}] \,.\, E[2\boldsymbol{\alpha}_i + \mathcal{S}_{ii}]$
- $= 2E[\boldsymbol{\alpha}_i]^2 + E[\boldsymbol{\alpha}_j \mathcal{S}_{ii}] +$
- $2E[\boldsymbol{\alpha}_i \boldsymbol{\alpha}_j] + 2E[\boldsymbol{\alpha}_j \boldsymbol{S}_{ii}] + 2E[\boldsymbol{\alpha}_j \boldsymbol{S}_{ii}] + 2E[\boldsymbol{\alpha}_i \boldsymbol{S}_{ij}] + 2E[\boldsymbol{S}_{ii} \boldsymbol{S}_{ij}]$
- $= 2E[\boldsymbol{\alpha}_i]^2 + E[\boldsymbol{\alpha}_j \mathcal{S}_{ii}]$
- $= Va + D1$

Correlation between lines and hybrids

$$r_{L,H} = \frac{\sigma_{LH}}{\sigma_L \ \sigma_H}$$

$$r_{L,H} = \frac{Va + D1}{\sqrt{(2Va + 4D1 + D2)(Va + Vd)}}$$

$$r_{L,H} = \frac{Va}{\sqrt{(2Va)(Va + Vd)}}$$

$$r_{L,H} = \frac{Va}{\sqrt{(2Va)(Va + \psi Va)}}$$

$$r_{L,H} = \frac{Va}{\sqrt{(2Va)Va(1 + \psi)}}$$

- Within population
- $F_2 = D1 = D2 = 0$
- $H = Vd$
- $\boldsymbol{\psi} = Vd \ / \ Va$
- $Vd = \boldsymbol{\psi} Va$

$$r_{L,H} = \frac{Va}{Va\sqrt{2\ (1 + \psi)}}$$

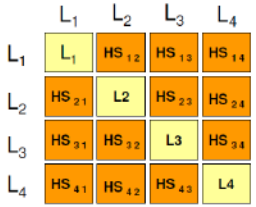
$$r_{L,H} = \frac{1}{\sqrt{2\ (1 + \psi)}}$$

- $Vd = 0; \boldsymbol{\psi} = 0$
- $Vd \ / \ Va = \boldsymbol{\psi} = ½$
- $Vd \ / \ Va = \boldsymbol{\psi} = 1$

- $r_{L,H} = 0.71$
- $r_{L,H} = 0.58$
- $r_{L,H} = 0.50$

Why mating designs?

- Estimate the components of variance
- Understand the genetic control
- Identify:
- *the best parents,*
- *populations structure (heterotic groups),*
- *testers, and*
- *the best combinations (hybrids)*
- Support decisions – *populations and breeding schemes*
- Over the years breeders have moved from full diallel to top cross – practical issues
- Ex. 49 lines, divided in two groups (34 and 15)



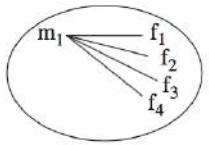
$$HS = n(n-1)/2$$

$$HS = \frac{49(48)}{2} = 1{,}176$$

	P1	P2
P3	F1(1,3)	F1(2,3)
P4	F1(1,4)	F1(2,4)

$$HS = na\ x\ n_b$$

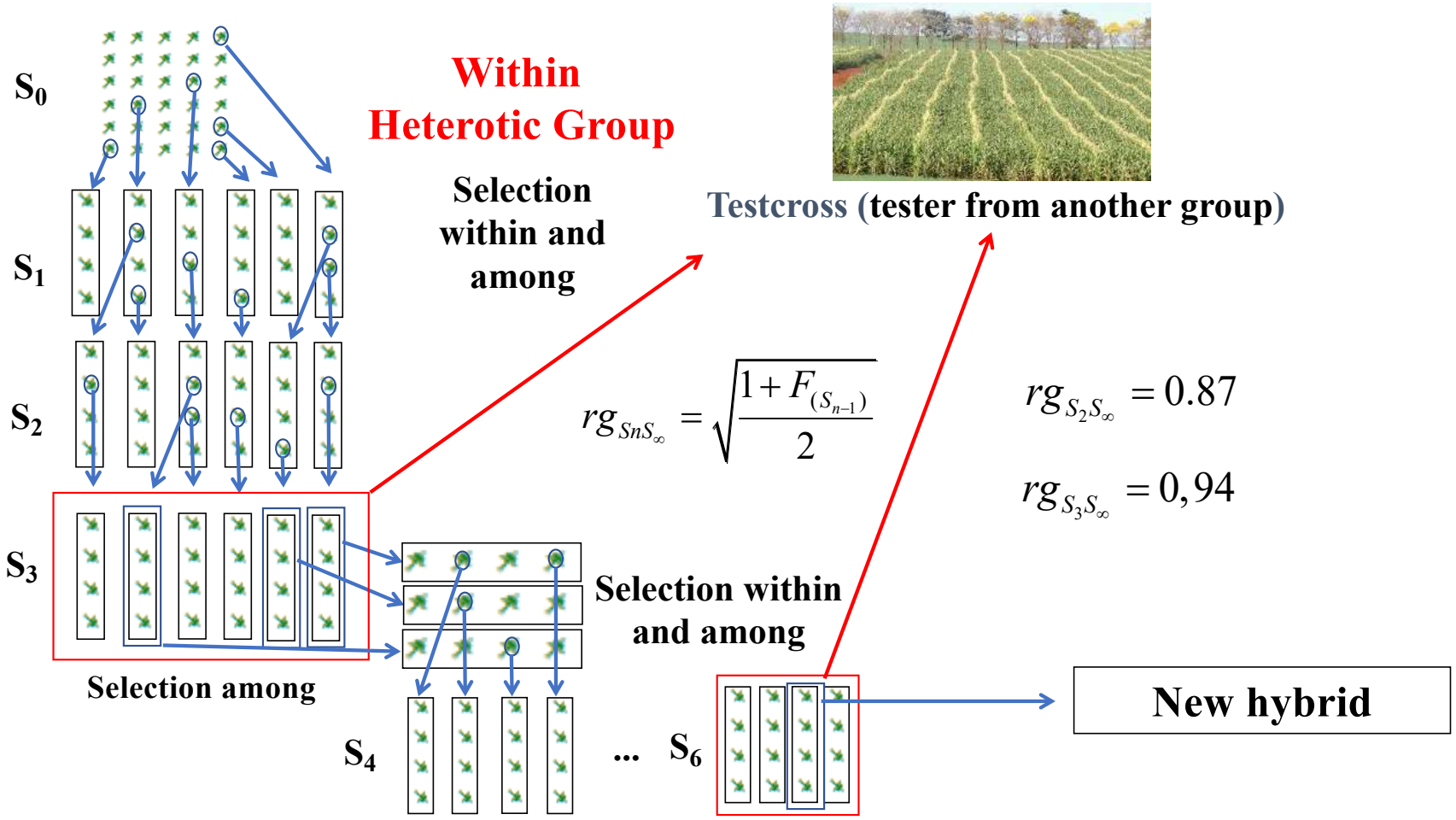
$$HS = 34\ x\ 15 = 510$$



$$HS = t_b n_a + t_a n_b$$

$$HS = 1x34 + 1x15 = 49$$

Obtaining lines – Early testcross



Early testcross

$$r_g = \frac{COV\ testcross(g, g')}{\sqrt{V_T g . V_T g'}}$$

$$V_T g = \frac{1}{2} pq[1 + Fg]\alpha_T^2$$

$$V_T g' = \frac{1}{2} pq[1 + Fg']\alpha_T^2$$

$$COV_{Tg,\,g'} = \frac{1}{2} pq[1 + Fg]\alpha_T^2$$

TABLE 4.2. Frequencies and testcross means of genotypes (F = inbreeding coefficient).

Population		Testcross progeny			Testcross
Genotype	Frequency	A_1A_1	A_1A_2	A_2A_2	mean
A_1A_1	p^2+pqF	p_T	q_T		$\mu_T+q\alpha_T$
A_1A_2	$2pq(1-F)$	$\frac{1}{2}p_T$	$\frac{1}{2}$	$\frac{1}{2}q_T$	$\mu_T+\frac{1}{2}(q-p)\alpha_T$
A_2A_2	q^2+pqF		p_T	q_T	$\mu_T - p\alpha_T$

$$r_g = \sqrt{\frac{1 + Fg}{1 + Fg'}} \qquad r_g = \sqrt{\frac{1 + Fg}{2}}$$

- *Normally, in the end we have Fg' = 1*

Early generation		Late generation	
Plant	Family	S6	Inbreds
S0	S1	0.71	0.71
S1	S2	0.87	0.87
S2	S3	0.94	0.94
S3	S4	0.98	0.97

Selecting for combining ability

- Increase the frequency of favorable alleles in lines
- Ideal tester:
- *Elite line = produce the new hybrid*
- *Single cross = produce a three-way cross hybrid*

Line	f(B)	Line x tester	CA
L1	p1	C1	CA1 = C1 – C..
L2	p2	C2	CA2 = C2 – C..
L3	p3	C3	CA3 = C3 – C..
…	…	…	…
L100	p100	C100	CA4 = C4 – C..
Mean	p	C..	

- $CA_i = (C_i - C..) = gi - \sum(p_i - p)\,\alpha_i^T$
- Lets consider two different lines
- $g_1 - \sum(p_1 - p)\,\alpha_1^T$
- $g_2 - \sum(p_2 - p)\,\alpha_2^T$
- $g_1 - g_2 = (p_1 - p)\,\alpha_1^T - (p_2 - p)\,\alpha_2^T$
- $g_1 - g_2 = (p_1 - p)\,\alpha_1^T - (p_2 - p)\,\alpha_2^T$
- $= (p_1 - p_2)\,\alpha^T$
- The difference is due to the frequency of favorable alleles

Choosing testers

- The best tester = correctly classify the lines
- Normally, it comes from the another heterotic group
- Should the tester be a elite or a poor line?
- Level of dominance and allele frequencies
- Consequences in breeding values

TABLE 4.2. Frequencies and testcross means of genotypes (F = inbreeding coefficient).

Population		Testcross progeny			Testcross
Genotype	Frequency	A_1A_1	A_1A_2	A_2A_2	mean
A_1A_1	p^2+pqF	p_T	q_T		$\mu_T+q\alpha_T$
A_1A_2	$2pq(1-F)$	$\frac{1}{2}p_T$	$\frac{1}{2}$	$\frac{1}{2}q_T$	$\mu_T+\frac{1}{2}(q-p)\alpha_T$
A_2A_2	q^2+pqF		p_T	q_T	$\mu_T-p\alpha_T$

$BV_i = (t_i - \bar{t})[a + (1-2r)d]$

$BV_i = (t_i - \bar{t})[a - 0.28]$

- genetic variability, and

$\sigma_T^2 = \frac{1}{2}pq(1+F)[a+(1-2r)d]^2$

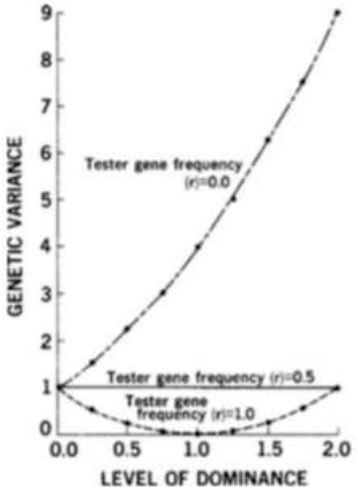
$\sigma_T^2 = pq[a-0.28]^2$

- expected gain (unrelated tester)

$\Delta_p = a + (1-2r)d$

$\Delta_p = a - 0.14$

$d = 0.7$
$r = 0.6$
$F = 1$



Genomic selection to predict Full Diallel and NCII