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

Response to Selection

Prof. Roberto Fritsche-Neto

roberto.neto@usp.br

Piracicaba, April 6th, 2018

Breeding value

- The breeding value (***a***) of a progeny ***i*** is the mean of its parents' plus the mendelian sampling

$$E(a_i) = \frac{1}{2}a_{p1} + \frac{1}{2}a_{p2} + m_i \qquad E(\bar{a}) = \frac{1}{2}a_{p1} + \frac{1}{2}a_{p2}$$

- It is key for deciding between possible breeding schemes

- Therefore, the expected response from the current to the next generation is entirely based on the genetic superiority of the selected parents

$$E(RS) = \frac{1}{2}\left(S_m + S_f\right)$$

- Based on the standard regression theory we have

$$b_{yx} = \frac{\sigma_{xy}}{x} = r_{xy}\frac{\sigma_y}{\sigma_x}$$

- Thus, we can predict the breeding value given an estimate which we will call the index value P
- P is the criteria for selection

$$\bar{\bar{a}}^* = \bar{a} + b_{aP}(P^* - \bar{P}) \qquad RS = \bar{\bar{a}}^* - \bar{a} = b_{aP}(P - \bar{P})$$

The Breeder's equation

- Using the deviation of the index values of selected values from the mean index value of all individuals in the population in standard deviations units

$$i = (P^* - \bar{P})/\sigma_P$$

$$(P^* - \bar{P}) = i.\sigma_P$$

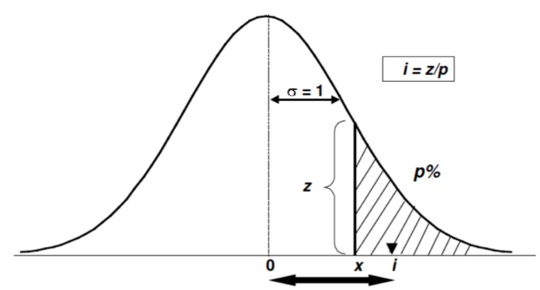
$$b_{aP} = r_{aP}\frac{\sigma_a}{\sigma_P}$$

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

$$RS = b_{aP}(P^* - \bar{P}) \qquad RS = b_{aP}.i.\sigma_P \qquad RS = r_{aP}\frac{\sigma_a}{\sigma_P}i.\sigma_P$$

$$RS = r_{aP}.\sigma_a.i \qquad RS = i.h.\sigma_a$$

- $i = z / p$ $\quad z = \frac{e^{-1/2x^2}}{\sqrt{2\pi}}$ $\quad p^* = \frac{s + 1/2}{n + \frac{s}{2n}}$
- p for small populations
- where ***s*** is the number selected and ***n*** is the population size



- On R, *i =dnorm(qnorm(p))/p*

- Response per unit time

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

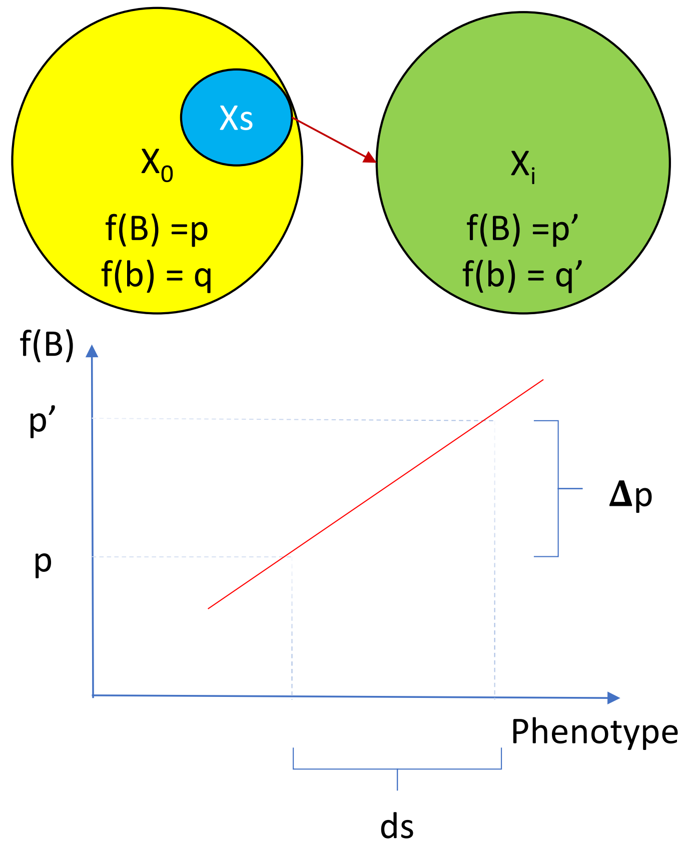
Effect of selection on allele frequencies

- $p + q = 1$
- $p' + q' = 1$
- $p' = p + \boldsymbol{\Delta} q$
- $q' = q - \boldsymbol{\Delta} p$
- $\boldsymbol{\Delta} p = \boldsymbol{\Delta} q$

- $(p' - p) = (u_s - u_0) b_{Pf(B)}$
- $\boldsymbol{\Delta} p = ds.\ b_{Pf(B)}$

- $b_{Pf(B)} = COV[P, f(B)] / V_P$

- $COV[P, f(B)] = COV[G + E, f(B)]$
- $= COV[G + E, f(B)]$
- $= COV[G, f(B)] + COV[E, f(B)]$
- $= COV[G, f(B)]$



Effect of selection on allele frequencies

- $COV(x,y) = \sum_i f_i x_i y_i - (\sum_i f_i x_i)(\sum_i f_i y_i)$

- $\sum_i f_i y_i = (p - q)a + 2pqd$

- $\sum_i f_i x_i = p^2(1) + 2pq(1/2) + q^2(0)$
- $= p^2 + pq$
- $= p^2 + p(1 - p)$
- $= p^2 + p - p^2 = p$

- $\sum_i f_i x_i y_i = p^2(1)(a) + 2pq(1/2)d + q^2(0)0$
- $= p^2 a + pqd$

- $COV(x,y) = p^2 a + pqd - p[(p - q)a + 2pqd]$
- $= p^2 a + pqd - p^2 a - pqa + 2p^2 d$
- $= pqd - pqa + 2p^2 d$
- $= pq[a + (1 - 2p)d]$
- $= pq[a + (q - p)d]$
- $= pq\boldsymbol{\alpha}$

Genotype	*f*	*VG*	*f(B)*
BB	p^2	a	1
Bb	2pq	d	1/2
bb	q^2	-a	0

- $\boldsymbol{\Delta} p = ds.\, b_{Pf(B)}$

- $b_{Pf(B)} = COV[P, f(B)] / V_P$

$$\Delta p = \frac{ds.p.q.\alpha}{\sigma_P^2}$$

$$\Delta p = \frac{i.p.q.\alpha}{\sigma_P}$$

$$(P^* - \bar{P}) = i.\sigma_P$$

Effect of selection on population mean

- $uc_o = u + (p - q)a + 2pqd$
- $uc_1 = u + (p' - q')a + 2p'q'd$

- $p' = p + \mathbf{\Delta}p$
- $q' = q - \mathbf{\Delta}p$
- $\mathbf{\Delta}p^2 \approx 0$

Genotype	C_0	C_1	VG
BB	p^2	P'^2	a
Bb	2pq	2p'q'	d
bb	q^2	q'^2	-a

- $uc_1 = u + (p' - q')a + 2p'q'd$
- $= u + ((p + \mathbf{\Delta}p) - (q - \mathbf{\Delta}p))a + 2d(p + \mathbf{\Delta}p)(q - \mathbf{\Delta}p)$
- $= u + (p + 2\mathbf{\Delta}p - q)a + 2pqd + 2qd\mathbf{\Delta}p - 2p\mathbf{\Delta}pd - 2\mathbf{\Delta}^2pd$
- $= u + (p - q)a + 2\mathbf{\Delta}pa + 2pqd + 2d\mathbf{\Delta}p(q - p)$
- $= u + (p - q)a + 2pqd + 2\mathbf{\Delta}pa + 2d\mathbf{\Delta}p(q - p)$
- $= u_0 + 2\mathbf{\Delta}pa + 2d\mathbf{\Delta}p(q - p)$
- $= u_0 + 2\mathbf{\Delta}p[a + d(q - p)]$
- $= u_0 + 2\mathbf{\Delta}p\boldsymbol{\alpha}$

- $\mathbf{\Delta}u = u_{c1} - u_{c0} = 2\mathbf{\Delta}p\boldsymbol{\alpha}$

$$\Delta p = \frac{i.p.q.\alpha}{\sigma_P}$$

$$\Delta u = 2\frac{i.p.q.\alpha}{\sigma_P}\alpha$$

$$\Delta u = i\frac{2pq\alpha^2}{\sigma_P} \qquad RS = i\frac{\sigma_a^2}{\sigma_P}$$

$$RS = i.h.\sigma_a$$

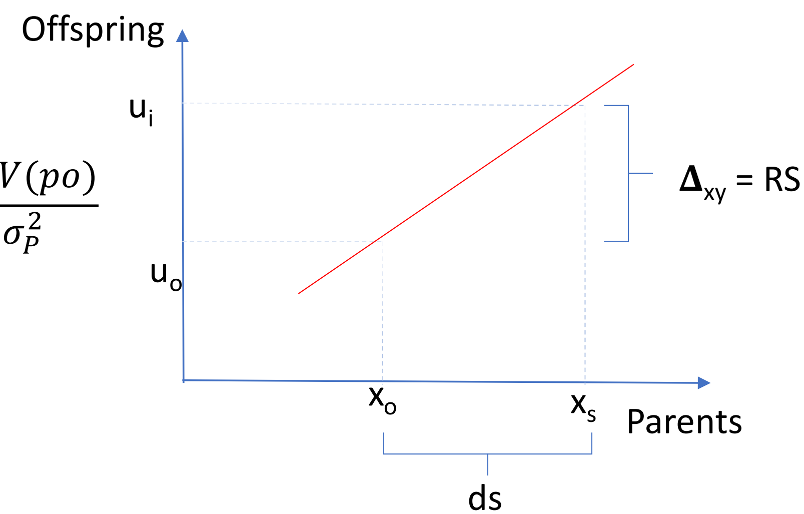
Response to Selection – parents to offspring

- Vx = Vp
- RS = ds.b_{xy} $\qquad i.\sigma_P = (P^* - \bar{P})$

$$RS = i.\sigma_P.\frac{COV(po)}{\sigma_P^2}$$

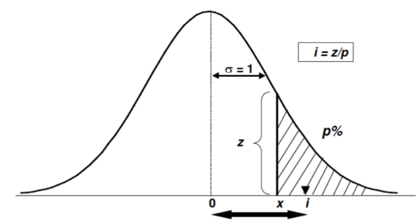
$$RS = \frac{i}{\sigma_P} cov_{po}$$

$$b_{po} = \frac{COV(po)}{\sigma_P^2}$$



- F = 0
- COV(p, o) = 2.f_{xy}.Va + u_{xy}Vd

Drift and Selection working together



- Frequencies after selection
- f(B)’ = p’ = p + Δp
- f(b)’ = q’ = q - Δp
- Frequencies after selection + drift (too small sample)
- E(p’) = p + Δp + *Sp*
- E(q’) = q - Δp - *Sp*
- E(*Sp*) = E(*Sq*) = 0

- *Sp* is the variance among samples $Sp = \frac{p'q'}{2Ne}$
- = (p + Δp) . (q - Δp) / 2Ne
- = (pq + Δp^2 - pΔp + qΔp) / 2Ne
- = (pq + Δp(q - p) / 2Ne
- = pq/2Ne + Δp(q - p)/2Ne
- drift per se + drift due to selection

$$\Delta p = i\frac{pq\alpha}{\sigma_P}$$

$$\sigma_A^2 = 2pq\alpha^2$$

$$DE = -2pqd$$

$$D1 = 2pq(q-p)\alpha d$$

$$\Delta F = \frac{1}{2Ne}$$

- uc_o = u + (p – q)a + 2pqd
- uc_1 = u + (p’ – q’)a + 2p’q’d

Drift and Selection working together

- $uc_1 = u + (p' - q')a + 2p'q'd$
- $= u + ((p + \mathbf{\Delta}p + Sp) - (q - \mathbf{\Delta}p - Sp))a + 2(p + \mathbf{\Delta}p + Sp)(q - \mathbf{\Delta}p + Sp)d$
- $= u + p + \mathbf{\Delta}p + Sp - q + \mathbf{\Delta}p + Sp)a + 2d(pq + p\mathbf{\Delta}p + pSp + q\mathbf{\Delta}p - \mathbf{\Delta}^2p - \mathbf{\Delta}pSp + qSp - \mathbf{\Delta}pSp - Sp^2)$
- $= u + (p - q)a + 2\mathbf{\Delta}pa + Spa + 2pqd + 2dp\mathbf{\Delta}p + 2dpSp + 2dq\mathbf{\Delta}p - 2d\mathbf{\Delta}pSp + 2dqSp - 2d\mathbf{\Delta}pSp - 2dSp^2)$

- Since $E(Sp) = 0$
- $= u + (p - q)a + 2\mathbf{\Delta}pa + 2pqd + 2dp\mathbf{\Delta}p + 2dq\mathbf{\Delta}p - 2dSp^2$
- $= 2\mathbf{\Delta}p[a + d\,(q - p)] - 2dSp^2$
- $= u_0 + 2\mathbf{\Delta}p\boldsymbol{\alpha} - 2\mathbf{\Delta}p(q - p)/2Ne - 2pqd/2Ne$

$$\Delta p = i\frac{pq\alpha}{\sigma_P}$$

$$ID = -2pqd$$

$$D1 = 2pq(q-p)\alpha d$$

$$\sigma_A^2 = 2pq\alpha^2$$

$$\Delta F = \frac{1}{2Ne}$$

$$\Delta x = u1 - u0 = i\frac{2pq\alpha^2}{\sigma_P} + \frac{i}{\sigma_P}\cdot\frac{2pq(q-p)\alpha d}{2Ne} - \frac{2pqd}{2Ne}$$

$$\Delta x = \frac{i\sigma_A^2}{\sigma_P} + \frac{iD1}{\sigma_P\ 2Ne} - \frac{ID}{2Ne}$$

$$\Delta x = \frac{i}{\sigma_P}\left(\sigma_A^2 + \frac{D1}{2Ne}\right) - \frac{ID}{2Ne}$$

$$RS = \frac{i}{\sigma_P}(\sigma_A^2 + \Delta FD1) - \Delta FDE$$

Indirect response to selection

- Pleiotropy and linked genes
- Hardly or costly to evaluate
- The second (***t***) trait should be high heritable and correlated with the first (***z***)
- $Vg_z = Va_z + Vd_z$
- $Vg_t = Va_t + Vd_t$
- $Va = 2pq[a + (q - p)d]^2$
- $Vd = (2pqd)^2$
- $COVg_{(z,t)} = COVa_{(z,t)} + COVd_{(z,t)}$

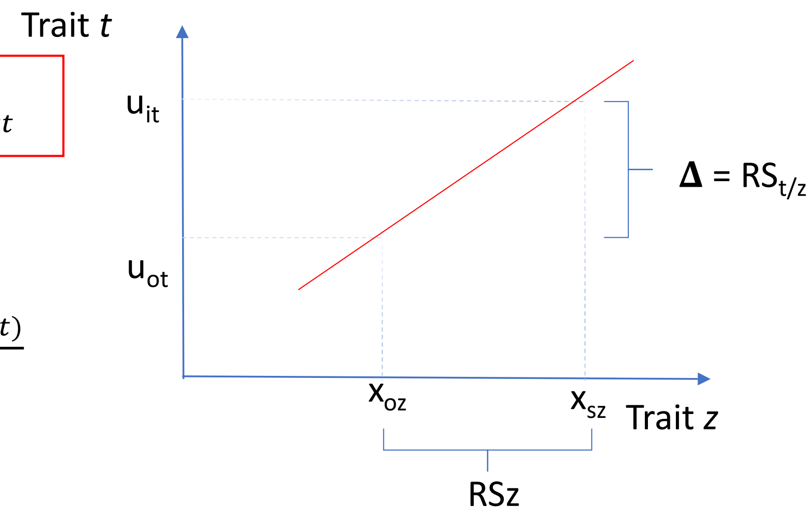
- F= 0
- $COV_g(z,t)_t = 2.f_{xy}.COVa_{(z,t)} + u_{xy}COVd_{(z,t)}$
- $COVa = 2pq[a_z + (q - p)d_z]\ [a_t + (q - p)d_t]$
- $COVd = 4p^2q^2d_t^2d_z^2$

$$RS(t,z) = RSz.b_{zt}$$

$$RSz = i\frac{\sigma_{Gz}^2}{\sigma_{Pz}}$$

$$b_{zt} = \frac{COV_{G(z,t)}}{\sigma_{Gz}^2}$$

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x\ \sigma_y}$$



$$RS(t,z) = i\frac{\sigma_{Gz}^2}{\sigma_{Pz}}\frac{COV_{G(z,t)}}{\sigma_{Gz}^2}$$

$$RS(t,z) = i\frac{COV_{G(z,t)}}{\sigma_{Pz}}$$

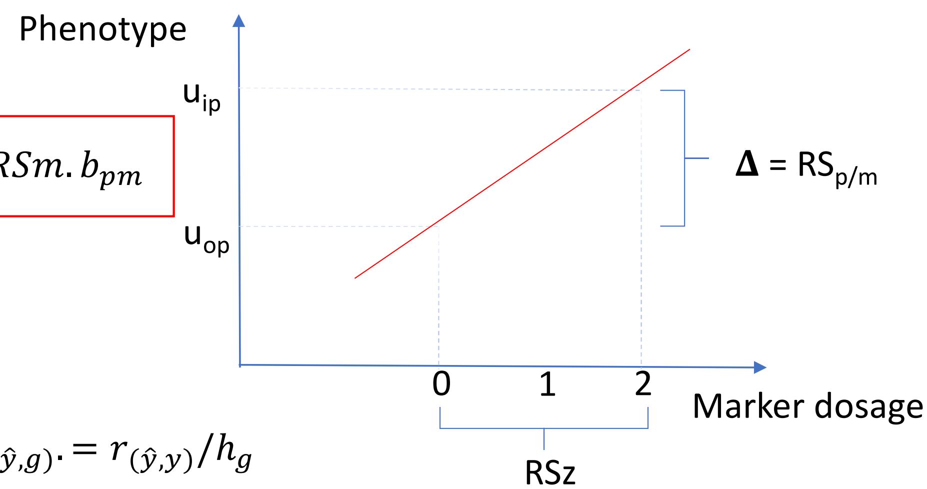
$$RS(t,z) = i\frac{r_{g(z,t)}\,\sigma_{gz}\sigma_{gt}}{\sigma_{Pz}}$$

$$RS(t,z) = ir_{g(z,t)}\,h_z\,\sigma_{gt}$$

Genomic Selection in(direct) response to selection

- The Breeder's equation

$$RS = \frac{i.r_a.\sigma_a}{T}$$



$$RS(p,m) = RSm.b_{pm}$$

- Predictive Ability vs. Accuracy

$$COV_{(\hat{y},y)} = COV_{(g_m,g+e)} = COV_{(gm,g)} + COV_{(gm,e)} = \sigma_g^2$$

$$V_{(\hat{y})} = V_{(g_m)} = \sigma_{g_m}^2 \qquad V_{(y)} = Vg + Ve = \sigma_g^2/h_g^2$$

$$r_{(\hat{y},y)} = \frac{COV_{(\hat{y},y)}}{\sigma_{\hat{y}}\ \sigma_y} = \frac{\sigma_g^2}{\sigma_{\hat{y}}\ .\sigma_g\ /h_g} = \frac{\sigma_g^2}{\sigma_{\hat{y}}\ \sigma_g}.h_g\ = r_{(\hat{y},g)}.h_g \qquad r_{(\hat{y},g)}. = r_{(\hat{y},y)}/h_g$$

- Response to selection in GS

$$RS(p,m) = i\frac{\sigma_{Gm}^2}{\sigma_{Pm}}.\frac{COV_{G(p,m)}}{\sigma_{Gm}^2} \qquad RS(p,m) = i\frac{COV_{G(p,m)}}{\sigma_{Pm}} \qquad RS(p,m) = i\frac{r_{g(p,m)}\,\sigma_{gp}\sigma_{gm}}{\sigma_{Pm}} \qquad RS(p,m) = ir_{g(p,m)}\,h_m\sigma_{gP}$$

$$b_{yx} = \frac{COV_{(x,y)}}{\sigma_x^2} \qquad r_{xy} = \frac{\sigma_{xy}}{\sigma_x\ \sigma_y} \qquad RS(p,m) = ir_{g(p,m)}\,\sigma_{gP}$$

Genomic Selection in(direct) response to selection

- Let's see why

$$V_{(y)} = Vg + Ve = \sigma_g^2 / h_g^2$$

- E(EBV) = b_{gy}.y
- V(cX) = c^2V(X)

$$\sigma_{EBV}^2 = b_{g,y}^2 \sigma_y^2$$

$$\sigma_y^2 = \frac{\sigma_{EBV}^2}{b_{gy}^2}$$

$$b_{g,y} = \frac{COV_{(g,y)}}{\sigma_y^2}$$

$$b_{g,y} = r_{g,y} = \frac{\sigma_{g,y}}{\sigma_g \ \sigma_y} = \frac{\sigma_g^2}{\sigma_g \ . \sigma_y} = \frac{\sigma_g}{\sigma_y} = h_g$$

$$\sigma_y^2 = \frac{\sigma_g^2}{h_g^2}$$

Is there a limit for RS?

- What is the RS over many cycles of selection?
- It is the sum of all cycles

- What is the maximum of RS? $RS = 2.Ne.i.r_{aP}.\sigma_a$

- Part of the observed variation in response is due to error when measuring response
- but part is due to drift variance
- The variance of total response because of drift can be predicted as:

$$RS = t.\frac{\sigma_a^2}{Ne}$$

- ***t*** is number of generations