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

Covariance between relatives

Prof. Roberto Fritsche-Neto

rfneto@agcenter.lsu.edu

Baton Rouge, Feb 6th, 2023

Definitions

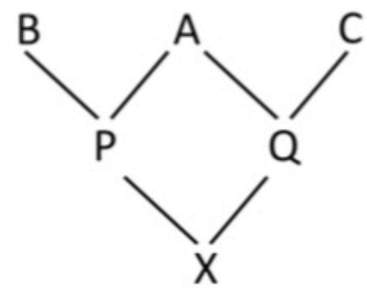
- Coefficient of kinship (f)
- Probability that two gametes taken at random from two individuals are identical by descent (IBD, ≡)
- Expresses the degree of relatedness between individuals - coefficient of parentage

- Coefficient of relationship (r)
- It is the additive genetic relationship between individuals
- This is the twice the coefficient of kinship
- $r = 2f$
- It is also equal the inbreeding coefficient of their progeny

- Aditive covariance between relatives
- The covariance between the breeding values
- $COV_{a(x,y)} = r_{xy}Va$
- It can actually be due to additive genetic effects, as well as dominance and epistatic effects
- In general the contribution of dominance and epistatic effects to the genetic covariance is low

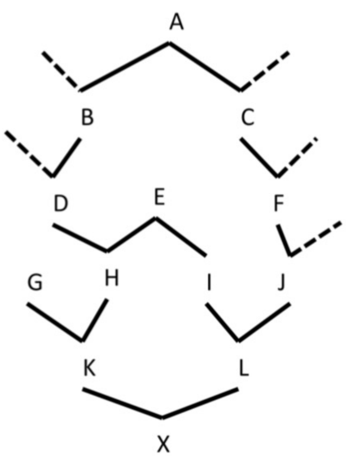
Calculating the inbreeding coefficient of X

- Lets consider one bi-allelic locus with two alleles A_1 and A_2
- We assume that A, the common ancestor of P and Q, is not inbred, thus its genotype is A_1A_2
- The probability that X receives A_1 from A via P, is the probability that A passes A_1 to P multiplied by the probability that P passes A_1 to X
- This probability is $1/2 \cdot 1/2 = 1/4$
- Now we need to know the probability that X receives A_1 from both P and Q
- $1/4 \cdot 1/4 = 1/16$
- We now know the probability that A_1 is IBD in X
- X could also be IBD by receiving two copies of A_2
- Probability IBD in X via either A_1 or A_2 is $1/16 + 1/16 = 2/16 = 1/8$
- $P(IBD) = 1/2^3 = 1/2^n$, where n is the number of common ancestral individuals
- However, if the parent A is inbred, the IBD increases and should be considered
- $(1/2)^n F_A$, where F_A is the inbreeding coefficient of the common ancestor
- Thus, IBD is the sum the two probabilities:
- $F_X = (1/2)^n + (1/2)^n F_A$
- $F_X = (1/2)^n(1+F_A)$



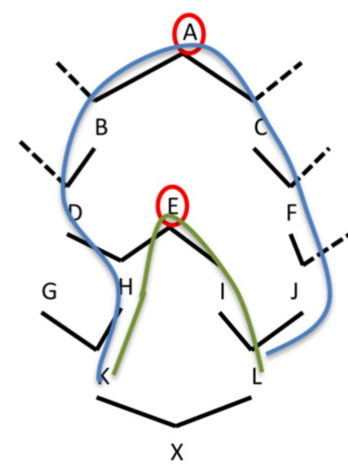
Calculating the inbreeding coefficient of X

- In more complex pedigrees, parents may be related to each other through more than one common ancestor, or from the same common ancestor, but through different paths
- The general formula is
- $F_X = (1/2)^n (1+F_A)$
- where ***n*** is the number of individuals in any path of relationship counting the parents of ***X*** and all individuals in the path which connects the parents to the common ancestor
- The summation is over all paths



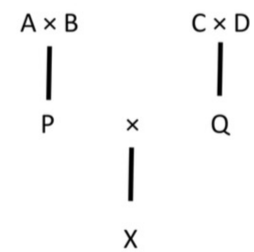
There are 2 common ancestors, A and E
There are 2 possible paths

Paths	n	F of common ancestor	Contribution of F_X
KHDB**A**CFJL	9	0	$(1/2)^9 = 0.002$
KH**E**IL	5	0	$(1/2)^5 = 0.031$
			Total$=$ 0.033



The coefficient of kinship

- Probability that two gametes taken at random (one from each individual carry alleles that are IBD
- The kinship (f) between two individuals is equal to the inbreeding coefficient of their progeny
- $F_X = f_{p1p2}$
- where X is the progeny and p_1 and p_2 are the parents



- Basic rules to estimate ***f***
- First: the ***f*** between P and Q is the mean of the four co-ancestries

$$f_{PQ} = \frac{1}{4} fAC + \frac{1}{4} fAD + \frac{1}{4} fBC + \frac{1}{4} fBD$$

- Second: the coefficient of kinship of an individual with itself f_{AA} is the inbreeding coefficient of progeny that would be produced by self-mating
- $f_{AA} = ½(1+F_A)$

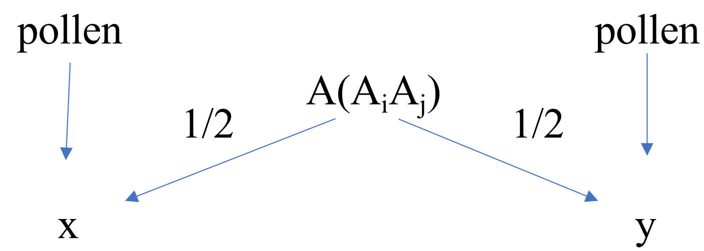
- Third: the coefficient of kinship between parent and offspring f_{PA} is the mean coefficient of kinship between A and both the parents of P, (A and B)
- $f_{PO} = ½(f_{AO} + f_{BO})$

Covariance between relatives

- $f_{xy} = ¼\, [P(x_i \equiv y_i) + P(x_j \equiv y_j) + P(x_i \equiv y_j) + P(x_j \equiv y_i)]$
- $u_{xy} = [P(x_i \equiv y_i;\, x_j \equiv y_j) + P(x_i \equiv y_j;\, x_j \equiv y_i)]$ – simultaneous events (the same genotype - dominance)

- **NON-INBRED RELATIVES**
- Half-sibs
- $F = 0$, thus, $F_A = 0$
- $f_{xy} = ¼\, [P(x \equiv y \equiv A_i) + P(x \equiv y \equiv A_j) + P(x \equiv A_i \equiv y \equiv A_j) + P(x \equiv A_j \equiv y \equiv A_i)]$
- If x and y are non-inbred, thus the last two parts are zero, because their parents are not inbred either
- $f_{xy} = ¼\, [1/4 + 1/4 + 0 + 0]$
- $f_{xy} = 1/8$
- Since $r = 2f$
- $r = 1/4$

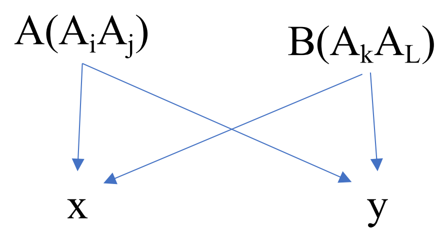
- $u_{xy} = 0$
- Probability of transmit the genotype – dominance effect



Covariance between relatives

- Full-sibs
- $F = 0$, thus, $F_A = F_B = 0$

- $f_{xy} = ¼ [P(x \equiv y \equiv A_i) + P(x \equiv y \equiv A_j) + P(x \equiv y \equiv A_k) + P(x \equiv y \equiv A_L) +$
- $P(x \equiv A_i \equiv y \equiv A_j) + P(x \equiv A_j \equiv y \equiv A_i) + P(x \equiv A_k \equiv y \equiv A_L) + P(x \equiv A_L \equiv y \equiv A_k)]$
- Since A and B are non-inbred, $P(A_i \equiv A_j) = 0$
- $f_{xy} = ¼ [1/4 + 1/4+ 1/4 + 1/4 + 0 + 0 + 0 + 0]$
- $f_{xy} = 1/4$
- Since $r = 2f$
- $r = ½$
- (½ . ½) . (½ . ½)
- $u_{xy} = [P(x \equiv y \equiv A_i; x \equiv y \equiv A_k) + P(x \equiv y \equiv A_i; x \equiv y \equiv A_L) +$
- $P(x \equiv y \equiv A_j; x \equiv y \equiv A_k) + P(x \equiv y \equiv A_j; x \equiv y \equiv A_L)]$
- $u_{xy} = [(0.25 \times 0.25) + (0.25 \times 0.25) + (0.25 \times 0.25) + (0.25 \times 0.25)]$
- $u_{xy} = 1/16 + 1/16 + 1/16 + 1/16$
- $u_{xy} = ¼$



Covariance between relatives

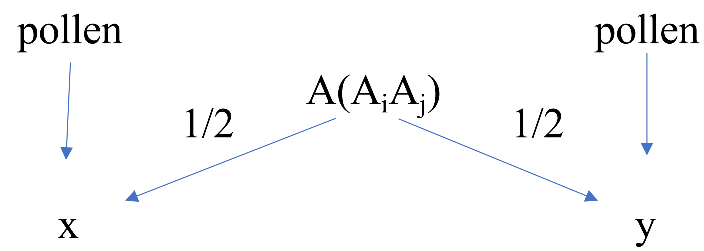
- **INBRED RELATIVES**

- Half-sibs
- $F_p \neq 0; P(A_i \equiv A_j) = P(A_k \equiv A_L) \neq 0$

- $f_{xy} = ¼\ [P(x \equiv y \equiv A_i) + P(x \equiv y \equiv A_j) + P(x \equiv A_i \equiv y \equiv A_j) + P(x \equiv A_j \equiv y \equiv A_i)]$
- these cases can be $A_i \equiv A_j$
- $f_{xy} = ¼\ [1/4 + 1/4 + F(1/4) + F(1/4)]$
- $f_{xy} = ¼\ [1/2 + F(1/2)]$
- $f_{xy} = 1/8[1 + F]$

- Since $r = 2f$
- $r = ¼[1+F]$

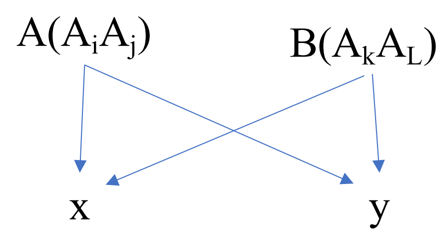
- $u_{xy} = 0$
- Probability of transmit the genotype – dominance effect



Covariance between relatives

- Full-sibs
- $F \neq 0$, thus, $F_A = F_B = 0$

- $f_{xy} = ¼\ [P(x \equiv y \equiv A_i) + P(x \equiv y \equiv A_j) + P(x \equiv y \equiv A_k) + P(x \equiv y \equiv A_L) +$
- $P(x \equiv A_i \equiv y \equiv A_j) + P(x \equiv A_j \equiv y \equiv A_i) + P(x \equiv A_k \equiv y \equiv A_L) + P(x \equiv A_L \equiv y \equiv A_k)]$
- Since A and B are non-inbred, $P(A_i \equiv A_j) = 0$
- $f_{xy} = ¼\ [1/4 + 1/4 + 1/4 + 1/4 + F(1/4) + F(1/4) + F(1/4) + F(1/4)]$
- $f_{xy} = ¼[1 + F]$
- Since $r = 2f$
- $r = ½[1 + F]$



- $(½\ .\ ½) + (½\ .\ ½\ )F$
- $u_{xy} = [P(x \equiv y \equiv A_i;\ x \equiv y \equiv A_k) + P(x \equiv y \equiv A_i;\ x \equiv y \equiv A_L) +$
- $P(x \equiv y \equiv A_j;\ x \equiv y \equiv A_k) + P(x \equiv y \equiv A_j;\ x \equiv y \equiv A_k)]$
- $u_{xy} = [(0.25 \text{ x } 0.25) + (0.25 \text{ x } 0.25) + (0.25 \text{ x } 0.25) + (0.25 \text{ x } 0.25)]$
- $u_{xy} = 1/16\ (1+F)^2 + 1/16(1+F)^2 + 1/16(1+F)^2 + 1/16(1+F)^2$
- $u_{xy} = ¼(1+F)^2$

Covariance between relatives

- $P(x \equiv y \equiv A_i) = P(x \equiv y \equiv A_k) + P(x \equiv A_i \equiv y \equiv A_j)$
- $= ½ . ½ + F(½ . ½)$
- $= ¼ + ¼F$
- $= ¼(1 + F)$

- $P(x \equiv y \equiv A_k) = P(x \equiv y \equiv A_k) + P(x \equiv A_k \equiv y \equiv A_L)$
- $= ½ . ½ + F(½ . ½)$
- $= ¼ + ¼F$
- $= ¼(1 + F)$

- $P(x \equiv y \equiv A_i; x \equiv y \equiv A_k) = P(x \equiv y \equiv A_i) . P(x \equiv y \equiv A_k)$
- $= ¼(1+F) . ¼(1+F)$
- $= 1/16(1+F)^2$

Why $r = 2f$?

- $x_{ij} = u + \boldsymbol{\alpha}_{ix} + \boldsymbol{\alpha}_{jx} + \mathcal{S}_{ijx}$
- $y_{ij} = u + \boldsymbol{\alpha}_{iy} + \boldsymbol{\alpha}_{jy} + \mathcal{S}_{ijy}$

- $E(x_{ij}) = E(y_{ij}) = u$
- $E(\boldsymbol{\alpha}_i) = \sum_i p_i \boldsymbol{\alpha}_i = 0$
- $E(\boldsymbol{\alpha}_j) = \sum_j p_j \boldsymbol{\alpha}_j = 0$
- $E(\mathcal{S}_{ij}) = \sum_{ij} p_i p_j \mathcal{S}_{ij} = 0$
- $E(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j) = E(\boldsymbol{\alpha}_i)\ E(\boldsymbol{\alpha}_j) = 0$
- $E(\boldsymbol{\alpha}_i, \mathcal{S}_{ij}) = E(\boldsymbol{\alpha}_i)\ E(\mathcal{S}_{ij}) = 0$
- $E(\boldsymbol{\alpha}_j, \mathcal{S}_{ij}) = E(\boldsymbol{\alpha}_i)\ E(\mathcal{S}_{ij}) = 0$

- $COV(x_{ij}, y_{ij}) = E[x_{ij}-E(x_{ij})] . E[y_{ij}-E(y_{ij})]$
- $= [u + \boldsymbol{\alpha}_{ix} + \boldsymbol{\alpha}_{jx} + \mathcal{S}_{ijx} - u] . [u + \boldsymbol{\alpha}_{iy} + \boldsymbol{\alpha}_{jy} + \mathcal{S}_{ijy} - u]$
- $= E(\boldsymbol{\alpha}_{ix}, \boldsymbol{\alpha}_{iy}) + E(\boldsymbol{\alpha}_{ix}, \boldsymbol{\alpha}_{jy}) + E(\boldsymbol{\alpha}_{ix}, \mathcal{S}_{ijy}) + E(\boldsymbol{\alpha}_{jx}, \boldsymbol{\alpha}_{iy}) + E(\boldsymbol{\alpha}_{jx}, \boldsymbol{\alpha}_{jy}) + \ldots + E(\mathcal{S}_{ijx}, \mathcal{S}_{ijy})$

- $E(\boldsymbol{\alpha}_{ix}, \mathcal{S}_{ijy}) = 0$
- Covariance between allele effect and genotype (all cases)

Why $r = 2f$?

- $E(\boldsymbol{\alpha}_{ix},\boldsymbol{\alpha}_{iy}) = \sum_i p_i \boldsymbol{\alpha}_i P(x_i \equiv y_i)\boldsymbol{\alpha}_i = \sum_i p_i \boldsymbol{\alpha}_i^2 P(x_i \equiv y_i) = 1/2Va.\ P(x_i \equiv y_i)$
- $E(\boldsymbol{\alpha}_{ix},\boldsymbol{\alpha}_{jy}) = \sum_i p_i \boldsymbol{\alpha}_i P(x_i \equiv y_j)\boldsymbol{\alpha}_i = \sum_i p_i \boldsymbol{\alpha}_i^2 P(x_i \equiv y_j) = 1/2Va.\ P(x_i \equiv y_j)$
- $E(\boldsymbol{\alpha}_{jx},\boldsymbol{\alpha}_{iy}) = \sum_j p_j \boldsymbol{\alpha}_j P(x_j \equiv y_j)\boldsymbol{\alpha}_j = \sum j_i p_j \boldsymbol{\alpha}_j^2 P(x_j \equiv y_j) = 1/2Va.\ P(x_j \equiv y_j)$
- $E(\boldsymbol{\alpha}_{jx},\boldsymbol{\alpha}_{jy}) = \sum_j p_j \boldsymbol{\alpha}_j P(x_j \equiv y_i)\boldsymbol{\alpha}_j = \sum j_i p_j \boldsymbol{\alpha}_j^2 P(x_j \equiv y_i) = 1/2Va.P(x_j \equiv y_i)$

- $E(\mathcal{S}_{ijx},\mathcal{S}_{ijy}) = \sum_{ij} p_i p_j \mathcal{S}_{ij}[P(x_j \equiv y_i, x_j \equiv y_j) + [P(x_i \equiv y_j, x_j \equiv y_i)]\mathcal{S}_{ij}$
- $= \sum_{ij} p_i p_j \mathcal{S}_{ij}^2[P(x_j \equiv y_i, x_j \equiv y_j) + [P(x_i \equiv y_j, x_j \equiv y_i)]$
- $= Vd.[P(x_j \equiv y_i, x_j \equiv y_j) + [P(x_i \equiv y_j, x_j \equiv y_i)]$

- $COV(x_{ij}, y_{ij}) = E[x_{ij}\text{-}E(x_{ij})]\ .\ E[y_{ij}\text{-}E(y_{ij})]$
- $= [u + \boldsymbol{\alpha}_{ix} + \boldsymbol{\alpha}_{jx} + \mathcal{S}_{ijx} - u]\ .\ [u + \boldsymbol{\alpha}_{iy} + \boldsymbol{\alpha}_{jy} + \mathcal{S}_{ijy} - u]$
- $= E(\boldsymbol{\alpha}_{ix},\boldsymbol{\alpha}_{iy}) + E(\boldsymbol{\alpha}_{ix},\boldsymbol{\alpha}_{jy}) + E(\boldsymbol{\alpha}_{ix},\mathcal{S}_{ijy}) + E(\boldsymbol{\alpha}_{jx},\boldsymbol{\alpha}_{iy}) + E(\boldsymbol{\alpha}_{jx},\boldsymbol{\alpha}_{jy}) + \ldots + E(\mathcal{S}_{ijx},\mathcal{S}_{ijy})$

- $COV_a(x_{ij}, y_{ij}) = ½Va[P(x_j \equiv y_i) + P(x_j \equiv y_j) + P(x_i \equiv y_j) + P(x_j \equiv y_i)] = ½Va[4fxy] = 2fVa$
- $COV_d(x_{ij}, y_{ij}) = ½Vd.[P(x_j \equiv y_i, x_j \equiv y_j) + [P(x_i \equiv y_j, x_j \equiv y_i)] = Vd[u_{xy}] = u_{xy}Vd$

- $COV_g(x_{ij}, y_{ij}) = 2.f_{xy}.Va + u_{xy}Vd$