Lecture 2:
Resemblance and relatedness

Bruce Walsh lecture notes
Introduction to Quantitative Genetics
SISG (Module 9), Seattle
20 – 22 July 2020

Heritability

- Central concept in quantitative genetics
- Fraction of phenotypic variance due to additive genetic values (Breeding values)
  - $h^2 = V_A/V_P$
  - This is called the narrow-sense heritability
  - Phenotypes (and hence V_P) can be directly measured
  - Breeding values (and hence V_A) must be estimated
- Estimates of V_A require known collections of relatives

Broad-sense heritability

- Narrow-sense heritability h^2 applies when outcrossing,
  - $h^2 = Var(A)/Var(P)$
  - = the fraction of all trait variation due to variation in breeding (additive genetic) values
- Broad-sense heritability H^2 applies when selecting among a series of pure lines
  - $H^2 = Var(G)/Var(P)$
  - = the fraction of all trait variation due to variation in Genotypic values

Defining H^2 for Plant Populations

Plant breeders often do not measure individual plants (especially with pure lines), but instead often measure a plot or a block of individuals.

This replication can result in inconsistent measures of H^2 even for otherwise identical populations.

Let z_{ijkl} denote the value of the l-th replicate in plot k of genotype i in environment j. We can decompose this value as

$$z_{ijkl} = G_i + E_j + GE_{ij} + p_{ijk} + e_{ijkl}$$

Effect of the k-th plot

deviations of individual plants within this plot

Suppose we replicate the genotype over e environments, with r plots (replicates) per environment, and n individuals per plot.

If we set our unit of measurement as the average over all plots, the phenotypic variance for the mean of line i becomes

$$\sigma^2(\bar{z}_i) = \sigma^2_G + \sigma^2_E + \frac{\sigma^2_{GE}}{e} + \frac{\sigma^2_p}{er} + \frac{\sigma^2_e}{ern}$$

Thus, V_P, and $H^2 = V_G/V_P$, depend on our choice of e, r, and n

In order to compare board-sense heritabilities we need to use a consistent design (same values of e, r, and n)

Key observations

- The amount of phenotypic resemblance among relatives for the trait provides an indication of the amount of genetic variation for the trait.
- If trait variation has a significant genetic basis, the closer the relatives, the more similar their appearance
- The covariance between the phenotypic value of relatives measures the strength of this similarity, with larger Cov = more similarity

PAUSE

- In some lines, a small fraction of chicken eggs are laid without a hard shell, but rather are coated with a soft membrane, and hence quickly broken.
- A poultry scientist trying to improve this trait might turn to environment change (e.g., add more calcium to their diet) or genetic improvement.
- For this trait, h^2 is very small, while H^2 is close to one. What do these observations suggest in terms of potential improvement strategies?

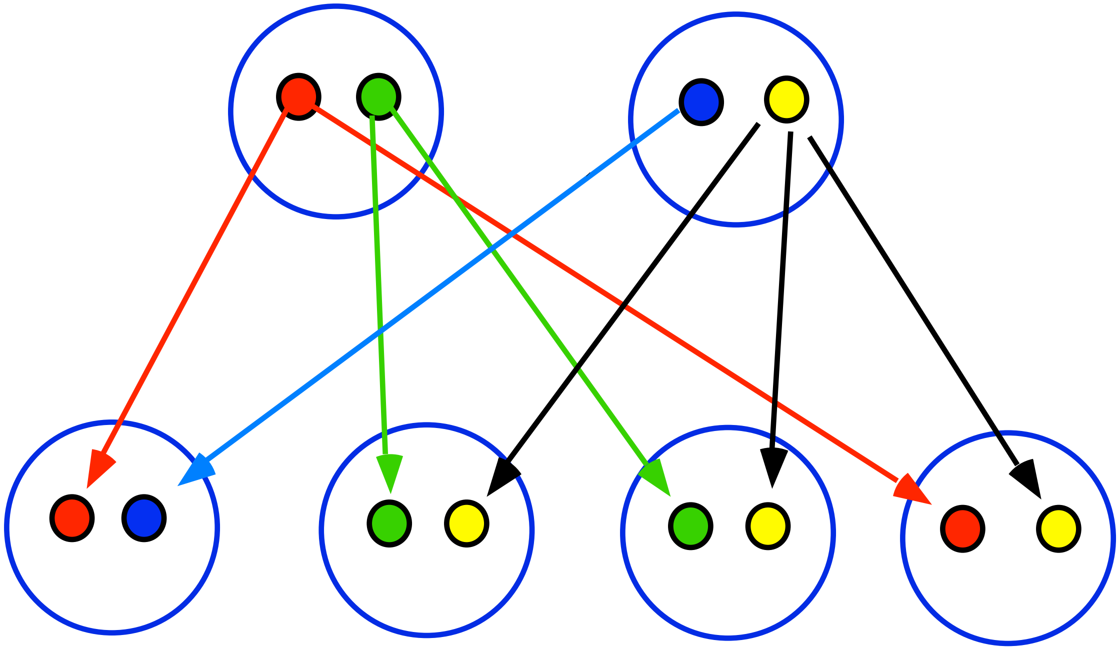
- Next pause at slide 20

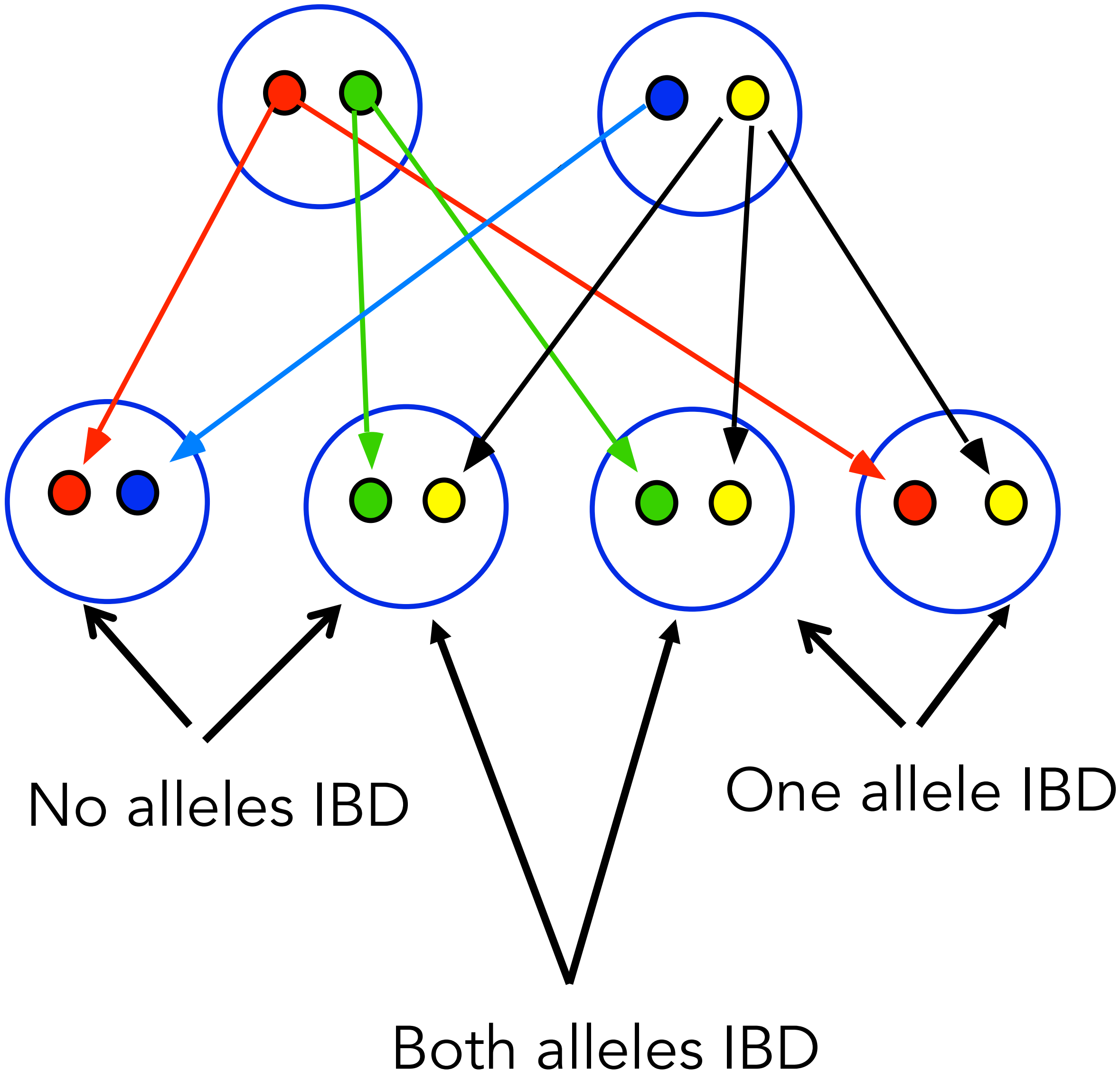
- H^2 close to one suggests most of the variation in this trait is genetic, so that environmental improvement (changes in management, such as using dietary supplements) is unlikely to impact the trait.
- h^2 close to zero suggests that there is little ADDITIVE variation, therefore only breeding schemes that exploit nonadditive variation (dominance, epitasis) are likely to make an improvement.

Genetic Covariance between relatives

Sharing alleles means having alleles that are identical by descent (IBD): both copies can be traced back to a single copy in a recent common ancestor.

Genetic covariances arise because two related individuals are more likely to share alleles than are two unrelated individuals.





No alleles IBD

One allele IBD

Both alleles IBD

Resemblance between relatives and variance components

- The phenotypic variance between relatives can be expressed in terms of genetic variance components
  - $Cov(z_x,z_y) = a_{xy}V_A + b_{xy}V_D$.
  - The weights a and b depend on the nature of the relatives x and y, and are measures of how often they are expected to share alleles identical by descent
  - These are critical in predicting selection response

Parent-offspring genetic covariance

$\mathrm{Cov}(G_p, G_o)$ --- Parents and offspring share
EXACTLY one allele IBD

Denote this common allele by A_1

$$G_p = A_p + D_p = \alpha_1 + \alpha_x + D_{1x}$$

$$G_o = A_o + D_o = \alpha_1 + \alpha_y + D_{1y}$$

IBD allele

Non-IBD alleles

$$
\begin{aligned}
Cov(G_o, G_p) &= Cov(\alpha_1 + \alpha_x + D_{1x}, \alpha_1 + \alpha_y + D_{1y} \\
&= Cov(\alpha_1, \alpha_1) + \cancel{Cov(\alpha_1, \alpha_y)} + \cancel{Cov(\alpha_1, D_{1y})} \\
&+ \cancel{Cov(\alpha_x, \alpha_1)} + \cancel{Cov(\alpha_x, \alpha_y)} + \cancel{Cov(\alpha_x, D_{1y})} \\
&+ \cancel{Cov(D_{1x}, \alpha_1)} + \cancel{Cov(D_{1x}, \alpha_y)} + \cancel{Cov(D_{1x}, D_{1y})}
\end{aligned}
$$

All blue covariance terms are zero.

- By construction, α and D are uncorrelated
  - By construction, α from non-IBD alleles are uncorrelated
  - By construction, D values are uncorrelated unless both alleles are IBD

$$Cov(\alpha_x, \alpha_y) = \begin{cases} 0 & \text{if } x \neq y, \quad \text{i.e., not IBD} \\ Var(A)/2 & \text{if } x = y, \quad \text{i.e., IBD} \end{cases}$$

$$Var(A) = Var(\alpha_1 + \alpha_2) = 2Var(\alpha_1)$$

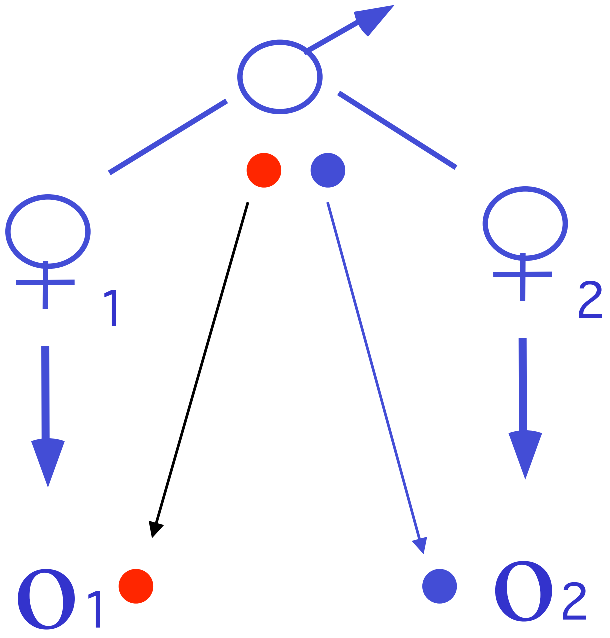
so that

$$Var(\alpha_1) = Cov(\alpha_1, \alpha_1) = Var(A)/2$$

Hence, relatives sharing one allele IBD have a genetic covariance of Var(A)/2

The resulting parent-offspring genetic covariance becomes $Cov(G_p, G_o) = Var(A)/2$

Half-sibs



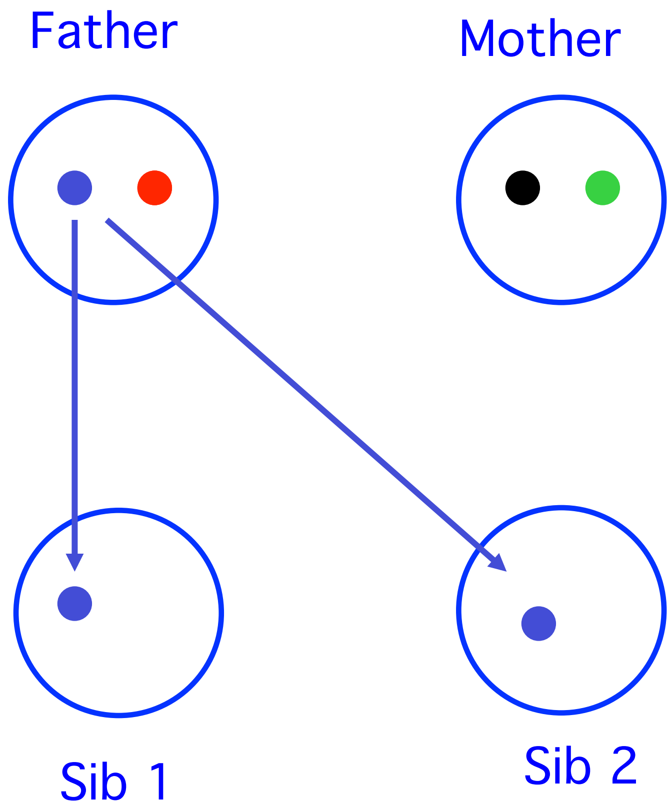
Each sib gets exactly one
allele from common father,
different alleles from the
different mothers

The half-sibs share no alleles IBD

- occurs with probability 1/2

Hence, the genetic covariance of half-sibs is just
(1/2)Var(A)/2 = Var(A)/4

Full-sibs

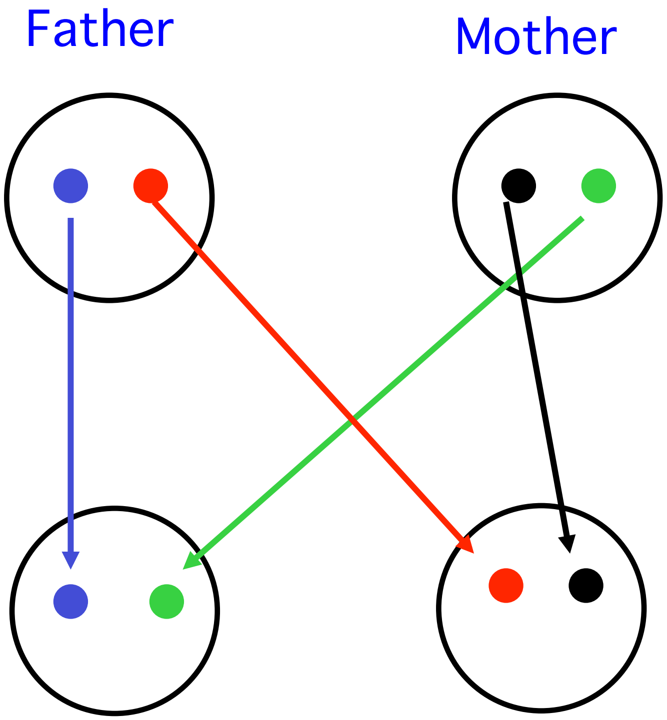


Each sib gets
exact one allele
from each parent

Prob(Allele from father IBD) = 1/2. Given the allele in parent one, prob = 1/2 that sib 2 gets same allele

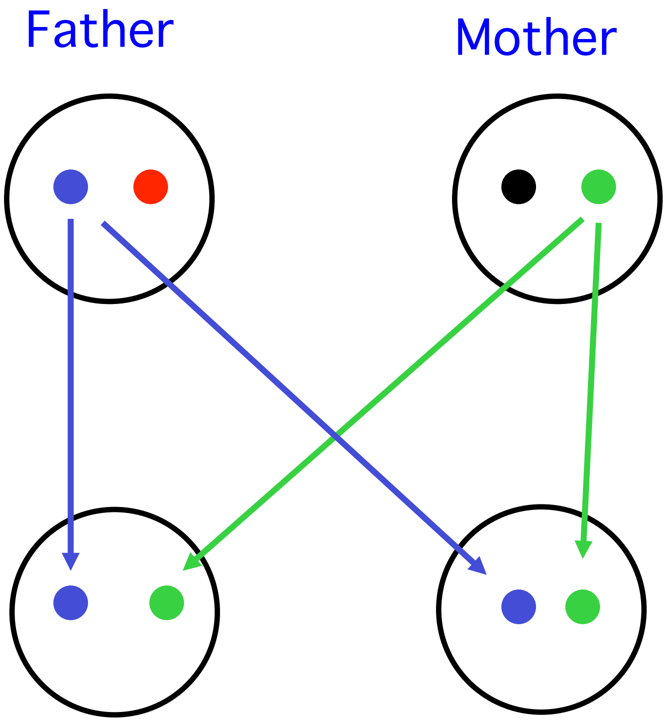
Prob(Allele from father not IBD) = 1/2. Given the allele in parent one, prob = 1/2 that sib 2 gets different allele

Full-sibs



Each sib gets
exact one allele
from each parent

Paternal allele not IBD [Prob = 1/2]
Maternal allele not IBD [Prob = 1/2]
Prob(sibs share 0 alleles IBD) = 1/2*1/2 = 1/4



Each sib gets
exact one allele
from each parent

Paternal allele IBD [Prob = 1/2]
Maternal allele IBD [Prob = 1/2]
Prob(sibs share 2 alleles IBD) = 1/2*1/2 = 1/4

Prob(share 1 allele IBD) = 1-Pr(0) - Pr(2) = 1/2

Resulting Genetic Covariance between full-sibs

IBD alleles	Probability	Contribution
0	1/4	0
1	1/2	Var(A)/2
2	1/4	Var(A) + Var(D)

Cov(Full-sibs) = Var(A)/2 + Var(D)/4

PAUSE

- Suppose the parent-offspring covariance for a trait is 12. What can you say about the additive and dominance variances?
- Suppose that the full-sib covariance for the same trait is 20. What can you say now about the dominance variance?

- Next pause at slide 28

- Cov(P, O) = Var(A)/2, hence
  - Var(A)/2 = 12, or Var(A) = 24
  - Can not say anything about dominance variance
- Cov(full sibs) = Var(A)/2 + Var(D)/4,
  - Hence 24/2 + Var(D)/4 = 20, or Var(D)/4 = 8, or Var(D) = 32
- Complication: Esp. in animals, full sibs can share a common family environmental variance, Var(Em), such as a common maternal effect. Hence,
  - Cov(full sibs) = Var(A)/2 + Var(D)/4 + Var(Em)
  - Thus, all that we can say is that Var(D)/4 + Var(Em) = 8
  - Hence, can only say that $\mathrm{Var(Em)} \leq 8$ or $\mathrm{Var(D)} \leq 32$

Genetic Covariances for General Relatives

Let r = (1/2)Prob(1 allele IBD) + Prob(2 alleles IBD)

Let u = Prob(both alleles IBD)

General genetic covariance between relatives
$Cov(G) = rVar(A) + uVar(D)$

When epistasis is present, additional terms appear
$r^2Var(AA) + ruVar(AD) + u^2Var(DD) + r^3Var(AAA) +$

More general relationships

- To obtain the expected covariance for any set of relatives, we normally need only compute r and u for that set of relatives
- With general inbreeding, becomes more complex (as three other terms, in addition to V_A and V_D arise)
- With crosses involving inbred and/or related parents, values for r and u are different from those presented above.

Coefficients of Coancestry

Suppose we pick a single allele each at random from two relatives. The probability that these are IBD is called Θ, the coefficient of coancestry. In terms of our previous notation, $2\Theta = r =$ the coeff on Var(A)

Θ_{xy} denotes the coefficient for relatives x and y

Consider an offspring z from a (hypothetical) cross of x and y. $\Theta_{xy} = f_z$, the inbreeding coefficient of z. Why? Because the offspring of x and y each get a randomly-chosen allele from each parent. The probability f_z that both alleles are IBD (the probability of inbreeding) is thus just Θ_{xy}.

θ and the coefficient on V_A

- The coefficient on the additive variance for the relatives x and y is just $2\theta_{xy}$.
- To see this,
  - let A_iA_j denote the two alleles in x and A_kA_l those in y.
  - Cov(breeding values) = $\Pr(A_i \text{ ibd } A_k)\,\mathrm{cov}(\alpha_i, \alpha_k) + \Pr(A_i \text{ ibd } A_l)\,\mathrm{cov}(\alpha_i, \alpha_l) + \Pr(A_j \text{ ibd } A_k)\,\mathrm{cov}(\alpha_j, \alpha_k) + \Pr(A_j \text{ ibd } A_l)\,\mathrm{cov}(\alpha_j, \alpha_l) = 4\,\theta_{xy}\mathrm{Var}(\alpha)$
  - Since $\mathrm{Var}(A) = 2\mathrm{Var}(\alpha)$, $\mathrm{Cov} = 2\,\theta_{xy}\mathrm{Var}(A)$

Θ_{xx} : The Coancestry of an individual with itself

Self x, what is the inbreeding coefficient of its offspring?

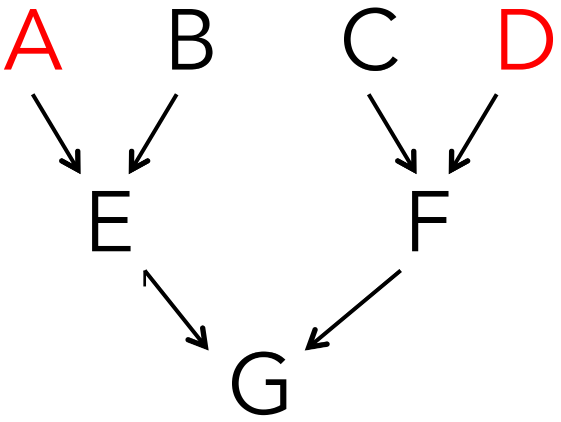
To compute Θ_{xx}, denote the two alleles in x by A_1 and A_2

	Draw A_1	Draw A_2
Draw A_1	IBD	f_x
Draw A_2	f_x	IBD

Hence, for a non-inbred individual, $\Theta_{xx} = 2/4 = 1/2$

If x is inbred, f_x = prob A_1 and A_2 IBD, $\boxed{\Theta_{xx} = (1+ f_x)/2}$

Example



Consider the following pedigree Suppose A and D are fully-inbred, and related, lines with $\theta_{AD} = 0.5$. Further, B and C are unrelated and outcrossed individuals

Individual	A	B	C	D
F_x	1	0	0	1
$\theta_{xx} = (1 + F_x)/2$	1	1/2	1/2	1

PAUSE

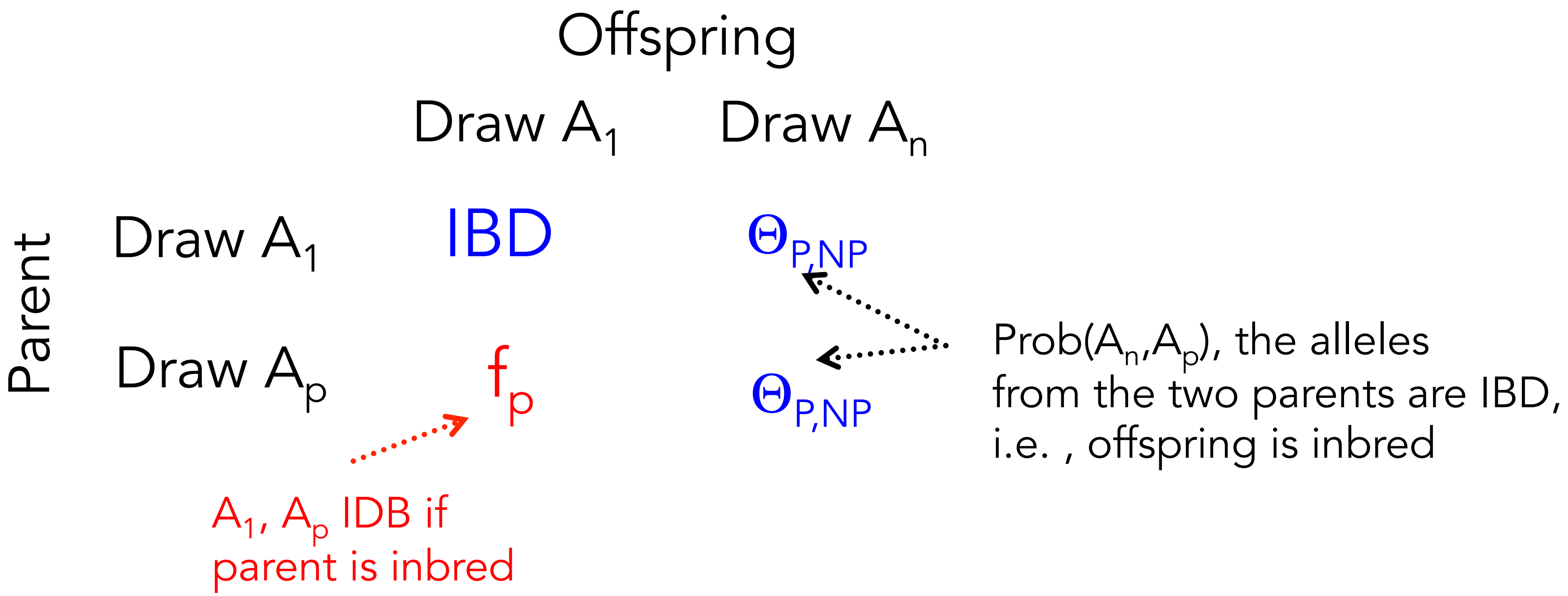
- Suppose Sam's COC is ¾.
  - What is Sam's level of inbreeding?
  - If we self Sam, what is the level of inbreeding in Sam's offspring?

- Next pause at slide 45

- $\frac{3}{4} = (1/2)(1+f_{Sam})$,
  - or $6/4 = 1+f_{Sam}$
  - Or $f_{Sam} = 6/4 - 1 = \frac{1}{2}$
- The COC between two individuals in the level of inbreeding in the offspring. Hence, we self Sam, the inbreeding in the offspring is simply Sam's COC or 3/4

The Parent-offspring Coancestry

Let A_1, A_n denote the two alleles in the offspring, where A_n is the allele from the nonfocal parent (NP), while A_1,A_p are the two alleles in the focal parent (P)



Offspring

Parent	Draw A_1	Draw A_n
Draw A_1	IBD	$\Theta_{P,NP}$
Draw A_p	f_p	$\Theta_{P,NP}$

Prob(A_n,A_p), the alleles from the two parents are IBD, i.e. , offspring is inbred

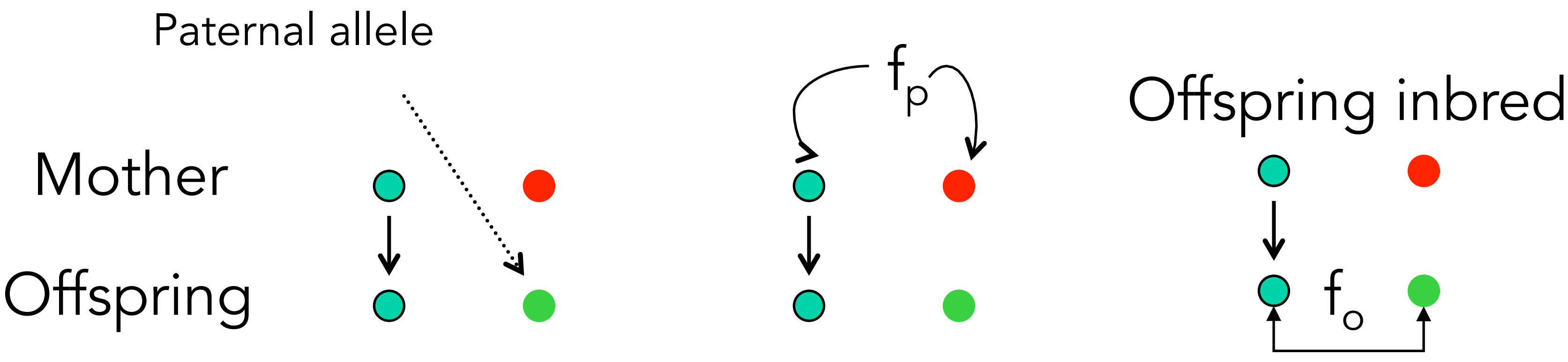
A_1, A_p IDB if parent is inbred

For a non-inbred individual, $\Theta_{P0} = 1/4$

General: $\boxed{\Theta_{PO} = (1 + f_p + 2\Theta_{P,NP})/4 = (1 + f_p + 2f_o)/4}$

Θ_{op} = Parent & Offspring

Parent inbred



$$\theta_{po} = \frac{1}{4}$$

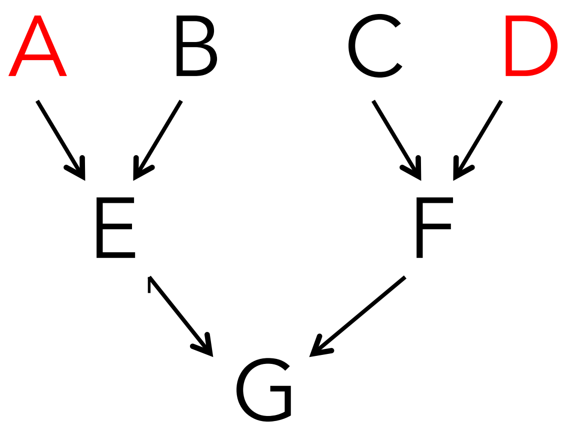
$$\theta_{po} = \frac{1+f_p}{4}$$

$$\theta_{po} = \frac{1+2f_o}{4}$$

1/2 = Prob random offspring allele from father. Prob = θ_{mf} = f_o that this allele is IBD to mother giving a contribution of $f_o/2$

$$\theta_{po} = \frac{1}{4}(1+f_p+2\theta_{mf})$$

This is just $2f_0$



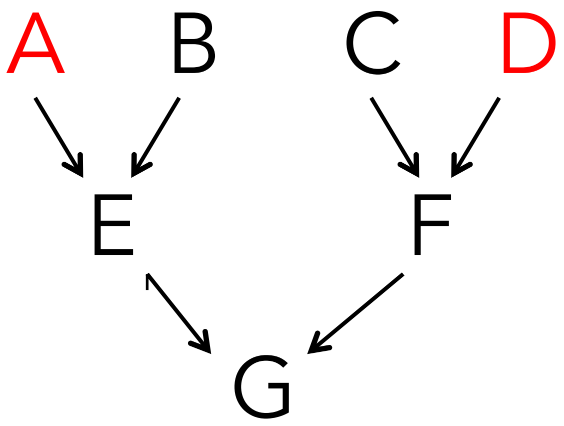
From before

$\theta_{AA} = \theta_{DD} = 1$; $\theta_{BB} = \theta_{CC} = 1/2$;
$\theta_{AD} = 1/2$,
$\theta_{AB} = \theta_{AC} = \theta_{BC} = \theta_{BD} = \theta_{CD} = 0$

Consider A - E (inbred parent - offspring)
$\theta_{AE} = (1+f_A)/4 = (1+1)/4 = 1/2$. Same value for θ_{DF}

Consider B - E (outbred parent - offspring)
$\theta_{BE} = (1+f_B)/4 = (1+0)/4 = 1/4$. Same value for θ_{CF}

Consider E - G (outbred parent - offspring)
$\theta_{EG} = (1+f_E)/4 = (1+0)/4 = 1/4$. Same value for θ_{FG}



From before

$\theta_{AA} = \theta_{DD} = 1$; $\theta_{BB} = \theta_{CC} = 1/2$;
$\theta_{AD} = 1/2$,
$\theta_{AB} = \theta_{AC} = \theta_{BC} = \theta_{BD} = \theta_{CD} = 0$

What about θ_{EF} ?

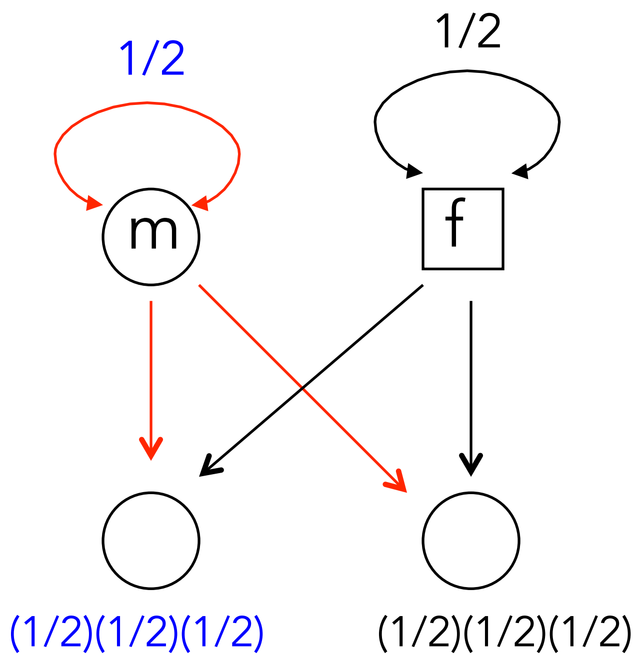
The randomly-chosen allele from E has equal chance of being from A or B. Likewise for F (from C or D)

Of these four possible combinations (A&C, A&D, B&C, B&D), only an allele from A and an allele from D have a chance of being IBD, which is $\theta_{AD} = 1/2$.

Hence, $\theta_{EF} = \theta_{AD}/4 = 1/8$

Full sibs (x and y) from parents m and f

$\Theta = 1/8 + 1/8 = 1/4$



1/2

1/2

m

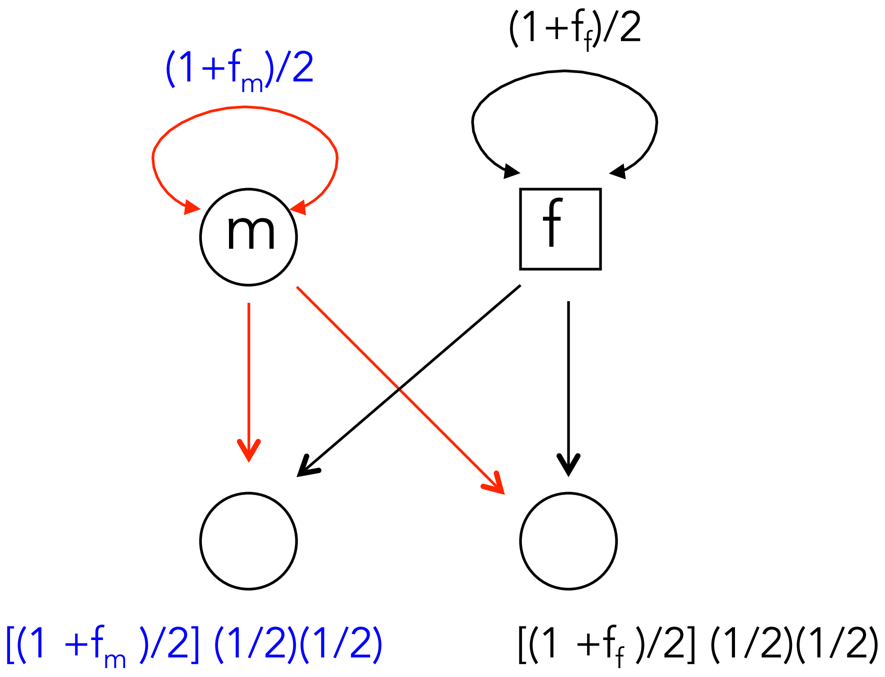
f

(1/2)(1/2)(1/2)

(1/2)(1/2)(1/2)

Unrelated, non-inbred
parents

$\Theta = (2 + f_m + f_f)/8$



$(1+f_m)/2$

$(1+f_f)/2$

m

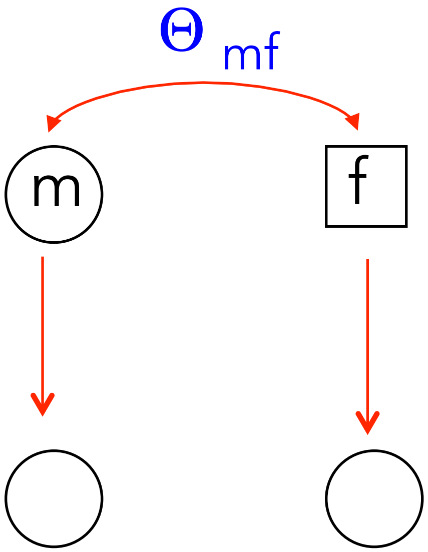
f

$[(1 + f_m)/2]\ (1/2)(1/2)$

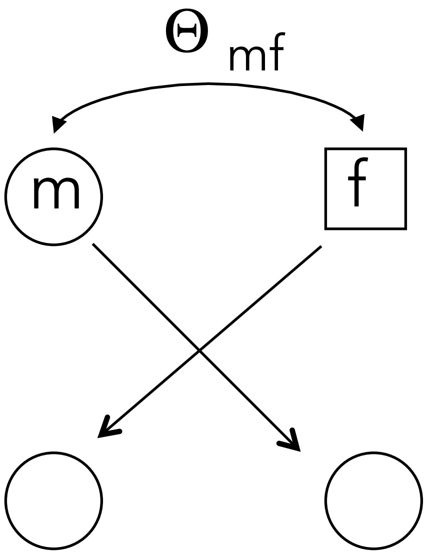
$[(1 + f_f)/2]\ (1/2)(1/2)$

Unrelated, inbred
parents

Full sibs (x and y) from parents m and f



Θ_{mf} (1/2)(1/2)



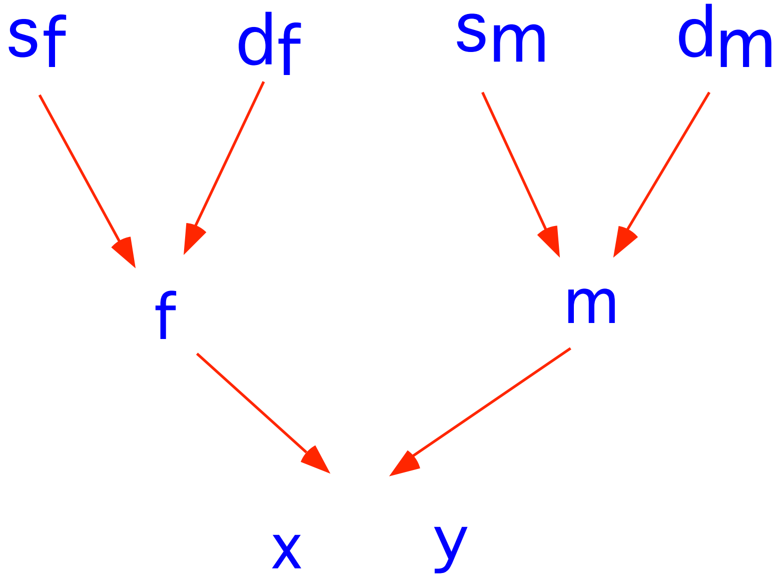
$\Theta_{mf}/4$

Parents inbred & related.
Two additional paths to add
to $\Theta = (2+f_m+f_f)/8$

This gives $\boxed{\Theta = (2+f_m+f_f+4\,\Theta_{mf})/8}$

Full sibs (x and y) from parents m and f

$$\Theta_{xy} = (2 + f_m + f_f + 4\Theta_{mf})/8$$



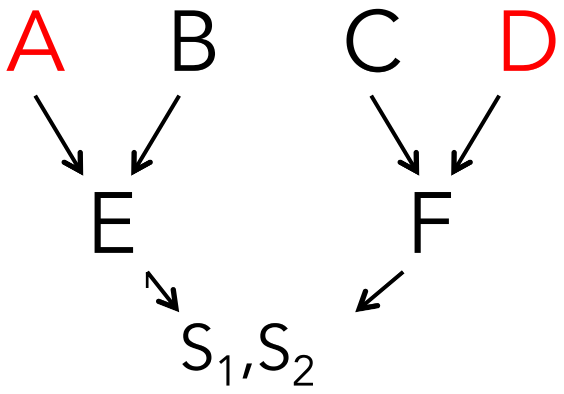
$$f_f = \Theta_{sf,df}$$

$$f_m = \Theta_{sm,dm}$$

Putting all this together gives

$$\boxed{\Theta_{xy} = (2 + \Theta_{sm,dm} + \Theta_{sf,df} + 4\Theta_{mf})/8}$$

Example



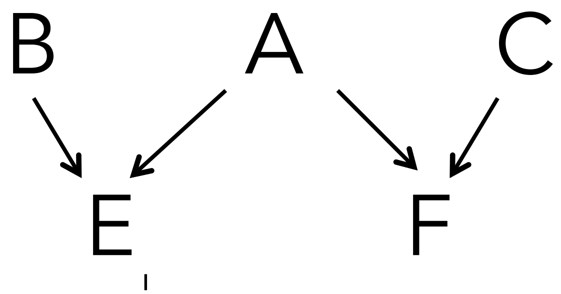
From before

$\theta_{AA} = \theta_{DD} = 1;\ \theta_{BB} = \theta_{CC} = 1/2;$
$\theta_{AD} = 1/2,\ \theta_{EF} = 1/8,$
$\theta_{AB} = \theta_{AC} = \theta_{BC} = \theta_{BD} = \theta_{CD} = 0$

$$\boxed{\Theta_{xy} = (2 + \Theta_{AB} + \Theta_{CD} + 4\Theta_{EF})/8}$$

$\theta_{S1S2} = (2 + 0 + 0 + 4[1/8])/8 = (4 + 1)/16 = 5/16$

Half-sibs



A is the common parent

- Using the same arguments as above,

$$\theta_{EF} = (\theta_{AA} + \theta_{AB} + \theta_{AC} + \theta_{BC})/4$$

$$= ([1 + f_A]/2 + \theta_{AB} + \theta_{AC} + \theta_{BC})/4$$

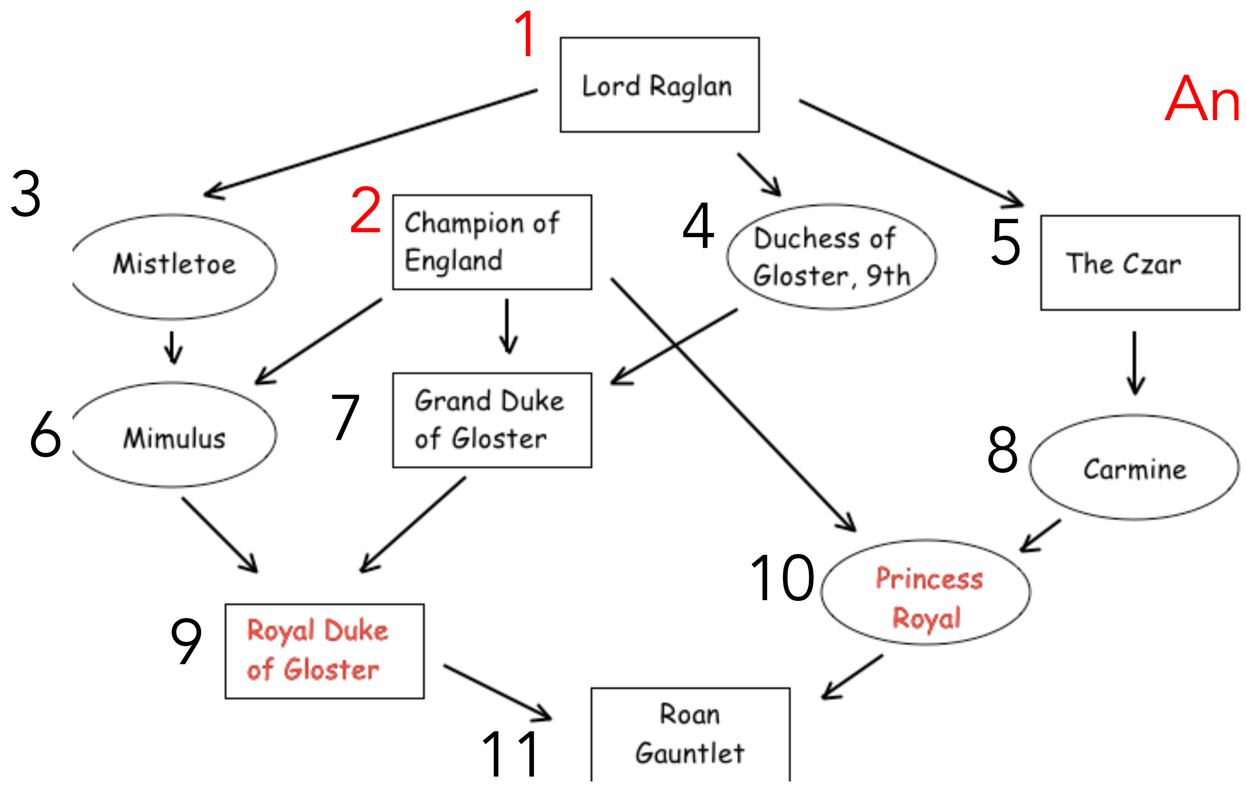
Hence, if B and C unrelated,

$$\theta_{EF} = (1 + f_A)/8$$

Computing θ_{xy} -- The Recursive Method

- There is a simple recursive method for generating the elements $A_{ij} = 2\,\theta_{ij}$ of a relationship matrix (used for BLUP selection). For ease of reading, we use the notation $A(i,j) = A_{ij}$
  - Basic idea is that the founding individuals of the pedigree are assumed to be unrelated and not inbred (although this can also be accommodated). These founders are assigned values of $A(i,i) = 1$.
  - Likewise, any unknown parent of any future individual is assumed to be unrelated to all others in the pedigree and not inbred, and they are also assigned a value of $A(i,i) = 1$.
  - Let S_i and D_i denote the sire and dam (father and mother) of individual i. For this offspring $A(i,i) = 1 + A(S_i, D_i)/2$
  - $A(i,j) = A(j,i) = [A(j,S_i) + A(j,D_i)]/2 = [A(i,S_j) + A(i,D_j)]/2$
  - The <u>recursive</u> (or <u>tabular</u>) method starts with the founding parents and then proceeds down the pedigree in a recursive fashion to fill out A for the desired pedigree.

Example

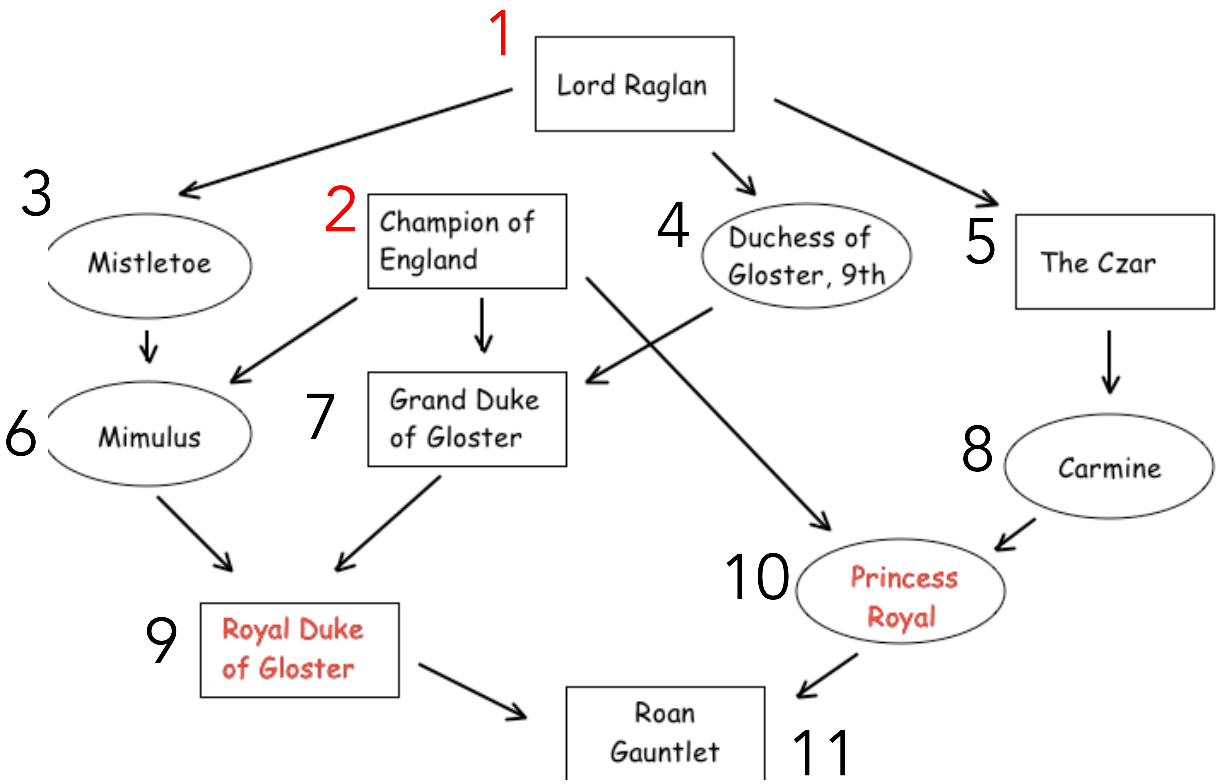


Ancestors are 1 & 2

$A(1,1) = A(2,2) = 1$
$A(1,2) = 0$

3, 4, 5, 8 all have unknown parents (only a single arrow to them)

3: $S_3 = 1$, $D_3 =$ Unknown, $A(3,3) = 1 + A(S_3,D_3)/2 = 1 + A(1,\text{unk})/2 = 1$
$A(1,3) = [A(1,S_3) + A(1,D_3)]/2 = [A(1,1) + A(1,\text{unk})]/2 = 1/2$.
Note also that $A(1,4) = A(1,5) = 1/2$, $A(4,4) = A(5,5) = 1$.
$A(3,4) = [A(3,S_4) + A(3,D_4)]/2 = [A(3,1) + A(3,\text{unk})]/2 = (1/2+0)/2 = 1/4$.
Same for $A(3,5) = 1/4$. 2 is unrelated to 3, 4, 5, giving $A(2,3) = A(2,4) = A(2,5) = 0$.



So far

	1	2	3	4	5
1	1	0	1/2	1/2	1/2
2	0	1	0	0	0
3	1/2	0	1	1/4	1/4
4	1/2	0	1/4	1	1/4
5	1/2	0	1/4	1/4	1

6: $S_6 = 2$, $D_6 = 3$. $A(6,6) = 1 + A(S_6, D_6)/2 = 1 + A(2,3)/2 = 1$

$A(6,1) = [A(1, S_6) + A(1, D_6)]/2 = [A(1,2) + A(1,3)]/2 = [0 + 1/2]/2 = 1/4$

$A(6,2) = [A(2, S_6) + A(2, D_6)]/2 = [A(2,2) + A(2,3)]/2 = [1+ 0]/2 = 1/2$

$A(6,3) = [A(3, S_6) + A(3, D_6)]/2 = [A(3,2) + A(3,3)]/2 = [0 + 1]/2 = 1/2$

$A(6,4) = [A(4, S_6) + A(4, D_6)]/2 = [A(4,2) + A(4,3)]/2 = [0 + 1/4]/2 = 1/8$

$A(6,5) = [A(5, S_6) + A(5, D_6)]/2 = [A(5,2) + A(5,3)]/2 = (0+1/4)/2 = 1/8$

7: $S_7 = 2$, $D_7 = 4$. $A(7,7) = 1 + A(S_7, D_7)/2 = 1 + A(2,4)/2 = 1 + 0/2 = 1$

$A(6,7) = [A(6, S_7) + A(6, D_7)]/2 = [A(6, 2) + A(6, 4)]/2 = (1/2 +1/8)/2 = 5/16$

8: $S_8 = 5$, $D_8 = \text{unk}$. $A(8,8) = 1 + A(S_8, D_8)/2 = 1 + A(5,\text{unk})/2 = 1$.

$A(6,8) = [A(6, S_8) + A(6, D_8)]/2 = [A(6, 5) + A(6, \text{unk})]/2 = (1/8)/2 = 1/16$

9: $S_9 = 7$, $D_9 = 6$. $A(9,9) = 1 + A(S_9, D_9)/2 = 1 + A(6,7)/2 = 1 + 5/32 = 1.156$ <- inbred!

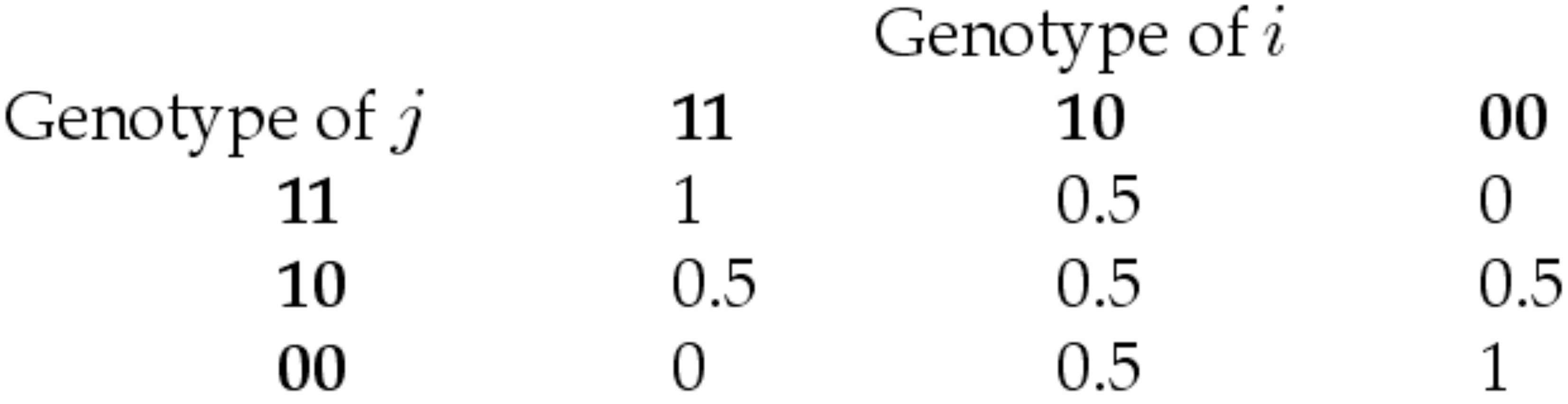
Actual relatedness versus expected values from pedigrees

Values for the coefficient of coancestry (θ) and the coefficient of fraternity (Δ) obtained from pedigrees are expected values. Due to random segregation of genes from parents, The actual value (or realization) can be different.

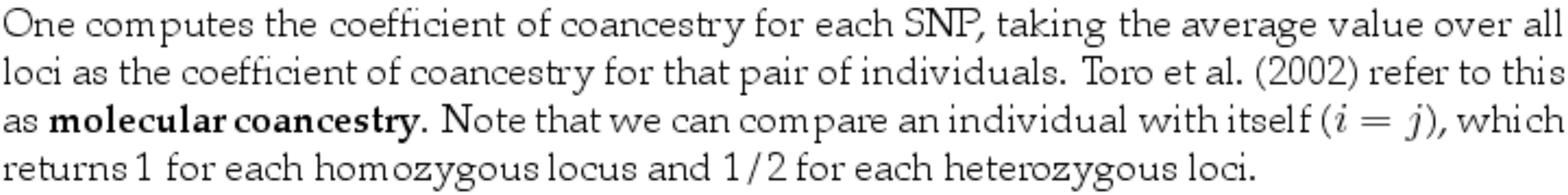
For example, we expect 2θ to be ½ for full subs. However, one pair of sibs may actually be more similar (0.6) and another less similar (say 0.35). On average, 2θ is ½ for pairs of full sibs, but if we knew the actual value of θ, we have more information. With sufficient dense genetic markers, we can estimate these relationships directly.

Genomic selection uses this extra information.

What about coefficient of coancestry θ ?



	Genotype of i		
Genotype of j	**11**	**10**	**00**
11	1	0.5	0
10	0.5	0.5	0.5
00	0	0.5	1



One computes the coefficient of coancestry for each SNP, taking the average value over all loci as the coefficient of coancestry for that pair of individuals. Toro et al. (2002) refer to this as **molecular coancestry**. Note that we can compare an individual with itself ($i = j$), which returns 1 for each homozygous locus and 1/2 for each heterozygous loci.

	Genotype of i		
Genotype of j	**11**	**10**	**00**
11	1	0.5	0
10	0.5	0.5	0.5
00	0	0.5	1

Indiv x:	00	00	10	10	00	10	11	00	11	00
Indiv y:	10	00	11	11	10	11	11	10	11	10
Locus-specific θ	0.5	1.0	0.5	0.5	0.5	0.5	1.0	0.5	1.0	0.5

Estimated θ is the average over all ten loci, = 0.65

PAUSE

– Key point: Θ, the coefficient of coancestry, is a central measure of the genetic relationship, with 2Θ giving the expected coefficient on contribution of Var(A) to the phenotypic correlation.
  - We showed how inbreeding and relatedness all **inflate Θ** over outbred and unrelated relatives,
  - With a known pedigree, the **expected value** of Θ for any set of pedigreed individuals can be computed
  - With dense (> 10,000 SNPS), we can use marker information to obtain the **realized**, as opposed to the expected (pedigree), value of Θ .
  - Next pause is slide 57

The coefficient of fraternity

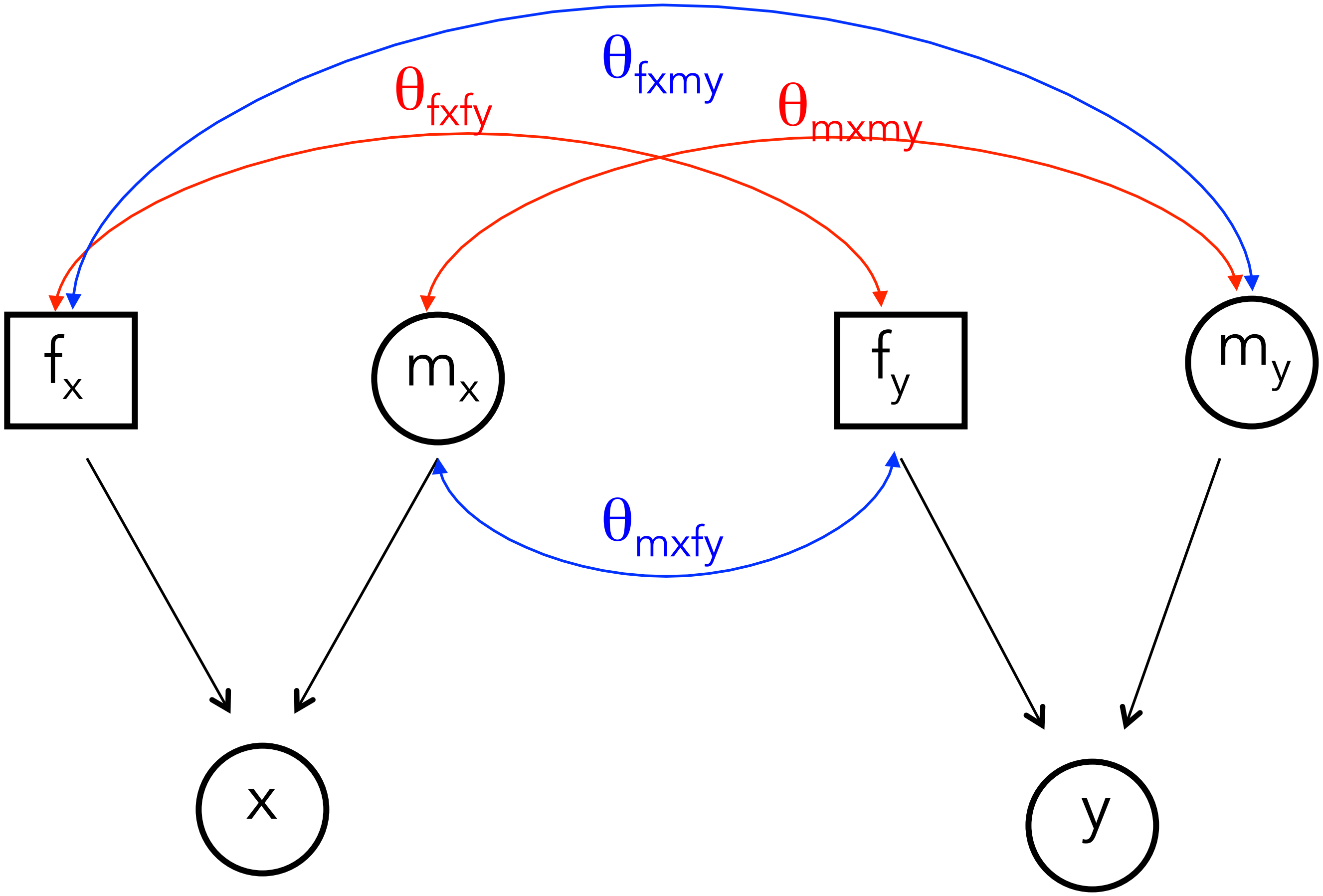
- While (twice) the coefficient of coancestry gives the weight on the additive variance for two relatives, a related measure of IDB status among relatives gives the weight on the dominance variance
- The probability that the two alleles in individual x are IBD to two alleles in individual y is denoted Δ_{xy}, and is called the coefficient of fraternity.
- This can be expressed as a function of the coefficients of coancestry for the parents of (mx and fx) of x and the parents (my and fy) of y.
  - $\Delta_{xy} = \theta_{mxmy}\theta_{fxfy} + \theta_{mxfy}\theta_{fxmy}$

The coefficient of fraternity (cont)

- x and y can have both alleles IBD if
  - The allele from the father (fx) of x and the father (fy) of y are IDB (probability θ_{fxfy}) AND the allele from the mother (mx) of x and the mother (my) of y are IDB (probability θ_{mxmy}) , or $\theta_{fxfy}\ \theta_{mxmy}$
  - OR the allele from the mother (mx) of x and the father (fy) of y are IDB (probability θ_{mxfy}) AND the allele from the father (fx) of x and the mother (my) of y are IDB (probability θ_{fxmy}) , or $\theta_{mxfy}\ \theta_{fxmy}$
  - Putting these together gives
    - $\Delta_{xy} = \theta_{mxmy}\theta_{fxfy} + \theta_{mxfy}\theta_{fxmy}$

Δ_{xy}, The Coefficient of Fraternity

Δ_{xy} = Prob(both alleles in x & y IBD)



$$\Delta_{xy} = \theta_{mxmy}\theta_{fxfy} + \theta_{mxfy}\theta_{fxmy}$$

Examples of Δ_{xy}: Full sibs

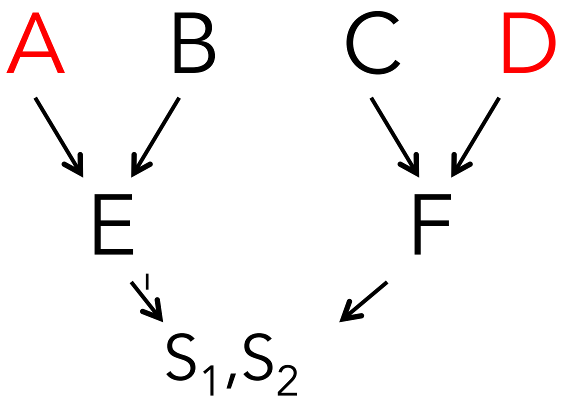
- Full sibs share same mon, dad
  - $m_x = m_y = m, \; f_x = f_y = f$
  - $\Delta_{xy} = \theta_{mxmy}\theta_{fxfy} + \theta_{mxfy}\theta_{fxmy} = \theta_{mm}\theta_{ff} + \theta_{mf}^2$
  - $\Delta_{xy} = (1+f_m)(1+f_f)/4 + \theta_{mf}^2$
- If parents unrelated, $\theta_{fm} = 0$, giving
  - $\Delta_{xy} = (1+f_m)(1+f_f)/4$
- If parents are unrelated and not inbred,
  - $\Delta_{xy} = 1/4$

Examples of Δ_{xy}: Half sibs

- Paternal half sibs share same dad, different moms
  - $f_x = f_y = f$; m_x and m_y
  - $\Delta_{xy} = \theta_{mxmy}\theta_{fxfy} + \theta_{mxfy}\theta_{fxmy} = \theta_{mxmy}\theta_{ff} + \theta_{mxf}\,\theta_{myf}$
  - $\Delta_{xy} = \theta_{mxmy}\,(1+f_m)/2 + \theta_{mxf}\,\theta_{myf}$
- If mothers are unrelated to each other and to the common father, $\theta_{mxmy} = \theta_{mxf} = \theta_{myf} = 0$, giving
  - $\Delta_{xy} = 0$

When is Δ non-zero?

- Since $\Delta_{xy} = \theta_{mxmy}\theta_{fxfy} + \theta_{mxfy}\theta_{fxmy}$
- A nonzero value for Δ requires either
  - That the fathers of both x and y are related AND the mothers of both x and y are related
  - OR that the father of x is related to the mother of y AND the mother of x is related to the father of y



From before

$\theta_{AA} = \theta_{DD} = 1$; $\theta_{BB} = \theta_{CC} = 1/2$;
$\theta_{AD} = 1/2$, $\theta_{EF} = 1/8$,
$\theta_{AB} = \theta_{AC} = \theta_{BC} = \theta_{BD} = \theta_{CD} = 0$

What is Δ for the full sibs (S_1 and S_2)?

$$\Delta_{xy} = \theta_{mxmy}\theta_{fxfy} + \theta_{mxfy}\theta_{fxmy} = \theta_{EE}\theta_{FF} + \theta_{EF}^2$$

Giving $\Delta_{xy} = \theta_{EE}\theta_{FF} + \theta_{EF}^2$

$$= (1/2)(1/2) + (1/8)^2$$

$$= 1/4 + 1/64 = 17/64 = 0.266$$

Δ_{xy} and the coefficient on V_D

- The coefficient on the dominance variance for the relatives x and y is just Δ_{xy}.
- To see this,
  - let A_iA_j denote the two alleles in x and A_kA_l those in y.
  - Suppose that alleles i and k come from the mothers of these two relatives and alleles j and l from their fathers.
  - Cov(dominance values) = $Pr(A_i \text{ ibd } A_{k;}\ A_j \text{ ibd } A_l)\ cov(\delta_{ij}, \delta_{kl}) + Pr(A_i \text{ ibd } A_{l;}\ A_j \text{ ibd } A_k) cov(\delta_{ij}, \delta_{kl})$
  - $= (\theta_{fxfy}\theta_{mxmy} + \theta_{mxfy}\theta_{jxmy})\ Var(D) = \Delta_{xy}\ Var(D)$

Estimating relationships using molecular data

With SNP data, treat identity in state (also called alike in state, AIS) as IBD

Suppose the genotypes of two individual at 10 SNPs are

Indiv x:	00	00	10	10	00	10	11	00	11	00
Indiv y:	10	00	11	11	10	11	11	10	11	10
		↑					↑		↑	

3/10 loci have $\Delta_{xy} = 1$, so average Δ_{xy} over all loci is $0.3 * 1 = 0.3$

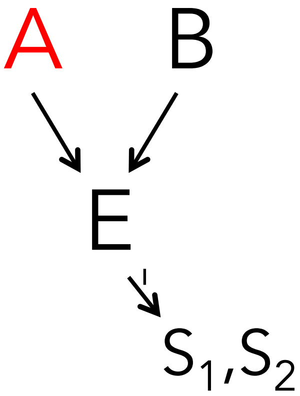
General Resemblance between relatives

$$2\theta_{xy} = r_{xy}, \qquad u_{xy} = \Delta_{xy}$$

$$Cov(G_x, G_y) = 2\theta_{xy} V_A + \Delta_{xy} V_D$$

$$\begin{aligned} Cov(G_x, G_y) = {} & 2\theta_{xy} V_A + \Delta_{xy} V_D \\ & + (2\theta_{xy})^2 V_{AA} + 2\theta_{xy} \Delta_{xy} V_{AD} + \Delta_{xy}^2 V_{DD} + \cdots \end{aligned}$$

Example



We found for full sibs S_1, S_2 that
$\theta = 5/16$, hence $2\theta = 5/8$; $\Delta = 17/64$

Expected genetic covariance between this sibs is

$$(5/8)\mathrm{Var}(A) + (17/64)\mathrm{Var}(D) + (5/8)^2\mathrm{Var}(AA) + (5/8)\,(17/64)\mathrm{Var}(AD) + (17/64)^2\mathrm{Var}(DD) + \cdots$$

PAUSE

- The coefficient of coancestry measures the fraction of alleles shared, and hence the contribution from additive variance (which is a function of the variation explained by single alleles)
- The coefficient of fraternity is the next level measure of relatedness, showing the fraction of diploid genotypes that are shared
  - It gives the weighting on the dominance variance
  - Can also be estimated from either pedigree of (dense) marker data.
  - Many, perhaps most, relationships can a co of fraternity of zero (as most have common relatives on both sides (mother and father) of the pedigree.
- Last slide is #60

Autotetraploids

- Peanut, Potato, alfalfa, soybeans all examples of crops with at least some autotetraploid lines
- With autotetraploid, four alleles per locus, with a parent passing along two alleles to an offspring
- As a result, a parent can pass along the <u>dominance contribution</u> in G to an offspring
- Further, now there are four variance components assocated with each locus

Genetic variances for autotetraploids

- G = A + D + T + Q
  - A (additive) and D (dominance, or digenic effects) as with diploids
  - T (trigenic effects) are the three-way interactions among alleles at a locus
  - Q (quadrigenic effects) are the four-way interactions at a locus
- Total genetic variance becomes
  - $V_G = V_A + V_D + V_T + V_Q$

Resemblance between autotetraploid relatives

Relatives	V_A	V_D	V_T	V_Q
Half-sibs	1/4	1/36		
Full-sibs	1/2	2/9	1/12	1/36
Parent -offspring	1/2	1/6		

Assumes unrelated, non-inbred parents