

Coefficients of Identity and Coancestry

Resemblance of Relative Pairs

- ▶ In our previous lectures we found that the phenotypic variance of a trait can theoretically be partitioned into a number of genetic and environmental components.
- ▶ How these components can be estimated in practice is of significant practical importance.
- ▶ A key principal for this is that genetic and environmental sources of variance contribute differentially to the resemblance between different types of relatives.
- ▶ In fact, much of population and quantitative genetics theory rests on the comparison of pairs of individuals - either on their genotypes or on their trait values, or both.

Resemblance of Relative Pairs

- ▶ Understanding the genetic covariance or correlation between relatives is essential in the determination of heritability of traits, as resemblance between relatives is a natural consequence of relatives inheriting copies of the same alleles.
- ▶ As is the case of the genetic variance, the genetic covariance between relatives can also be partitioned into components attributable to additive, dominance, and various epistatic effects.
- ▶ Each of these variance component terms for the decomposition of genetic covariances of a particular relative pair type are weighted by a coefficient that is calculated from the joint distribution of the alleles shared by the pair.
- ▶ Today we will focus on defining and calculating coefficients of relatedness for relative pairs.

Measures of Relatedness

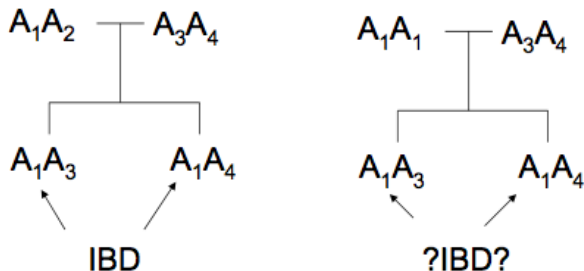
- ▶ A number of relatedness measures have been developed for population genetics.
- ▶ We will discuss some of the measures that play a central role in quantitative-genetic formulations.
- ▶ Measures of relatedness generally share two essential features.
- ▶ First, relatedness can only be defined relative to some specific frame of reference.
- ▶ All members of a species or population share a common ancestor at some point in time in a phylogeny.
- ▶ We will generally consider recent ancestry as opposed to ancient common ancestry.
- ▶ For example, if no further genealogical information is observed for the top generation of a pedigree, these individuals will be considered as the base and will be treated as unrelated individuals in the population.

Measures of Relatedness

- ▶ Second, measures of relatedness are based on the concept of **identity by descent (IBD)**.
- ▶ Alleles that are **IBD** are exact copies of an ancestral allele.
- ▶ The distinction between **identical by state (IBS)** and **IBD** is crucial.
- ▶ Alleles that have identical nucleotide sequences but have descended from different ancestors in the reference population are IBS but not IBD
- ▶ On the other hand, alleles that are IBD are necessarily IBS provided there is no mutation of the inherited allele.

Pedigree IBD Sharing Example 1

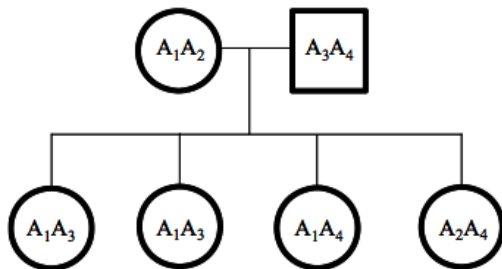
- Consider the two pedigrees below. Which alleles are shared IBD for the siblings in the two pedigrees?



- For the pedigree on the left, the A_1 allele is shared IBD by the siblings. For the pedigree on the right, with probability $\frac{1}{2}$ the A_1 allele is shared IBD by the siblings.

Pedigree IBD Sharing Example 2

- ▶ For the following nuclear family, there are six pairs of siblings. For each pair, how many alleles are shared IBD?



Identity States

- ▶ The IBD status of a single locus for an individual can be represented by F and $1 - F$, where F is the inbreeding coefficient):
- ▶ A complete description of the IBD status among the four alleles alleles carried by two individuals at a locus requires 15 identity states















Identity States

- ▶ Consider a single locus in two individuals. There are four alleles involved. We can describe IBD status using the following schematic. Start with a grid of four nodes:

	maternal	paternal
individual 1's alleles	•	•
individual 2's alleles	•	•

- ▶ A pair of nodes is connected with a line if the alleles are IBD. There are 15 possible configurations of IBD (next slide). We do not usually care about the distinction between maternal and paternal alleles, so the 15 possible configurations of IBD reduce to nine identity states.

Jacquard's (1970) 9 Condensed Coefficients of Identity

Probability	Identity state		Maternal	Paternal
Δ_1		A's alleles	•	•
		B's alleles	•	•
Δ_2				
Δ_3				
Δ_4				
Δ_5				
Δ_6				
Δ_7				
Δ_8				
Δ_9				

Jacquard's (1970) 9 Condensed Coefficients of Identity

- ▶ Δ_i is the probability the two individuals have IBD status of state i .
- ▶ For a given pair of individuals, Δ_1 through Δ_9 are the probabilities associated with each of the nine states.
- ▶ The first six states only occur if there is inbreeding, and in large, random-mating outbred populations, the probability of the first six states is essentially 0.
- ▶ That is, Δ_1 through Δ_6 have values of 0, and Δ_7 , Δ_8 , and Δ_9 are the probabilities that are of interest for relative pairs in outbred populations.
- ▶ Note that Δ_7 can be characterized as the probability that two individuals share both of their alleles IBD. Δ_7 is also called the **coefficient of fraternity**.

Coefficients of Coancestry

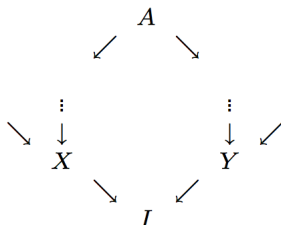
- ▶ Now consider two alleles at a locus that are drawn at random from two individuals i and j
- ▶ The probability that the two alleles are identical by descent (IBD) is defined to be the **coefficient of coancestry** or **kinship coefficient** and is often represented as θ_{ij} .
- ▶ θ_{ij} can be written in terms of the condensed coefficients of identity.
- ▶ Take a moment and write the kinship coefficient in terms of the condensed coefficients of identity

Coefficients of Coancestry

$$\theta_{ij} = \Delta_1 + \frac{1}{2}(\Delta_3 + \Delta_5 + \Delta_7) + \frac{1}{4}\Delta_8$$

Path Counting

- ▶ Algorithms to obtain kinship coefficients often use a technique called "path counting"



- ▶ To get the kinship coefficient for I 's parents, we would identify the path linking the parents of I to their common ancestor(s).

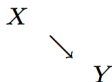
Path Counting

- ▶ If the parents X, Y of an individual I have ancestor A in common, and if there are n individuals (including X, Y, I) in the path linking the parents through A , then the inbreeding coefficient of I , or the coancestry of X and Y , is

$$F_I = \theta_{XY} = \left(\frac{1}{2}\right)^n (1 + F_A)$$

- ▶ If there are several ancestors, this expression is summed over all the ancestors.

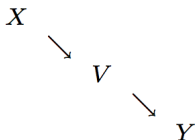
Path Counting: Parent-Offspring



- ▶ The common ancestor of parent X and child Y is X . The path linking X, Y to their common ancestor is YX and this has $n = 2$ individuals. Therefore

$$\theta_{XY} = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

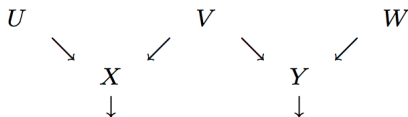
Path Counting: Grandparent-Grandchild



- ▶ The common ancestor of grandparent X and grandchild Y is X . The path linking X , Y to their common ancestor is YVX and this has $n = 3$ individuals. Therefore

$$\theta_{XY} = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

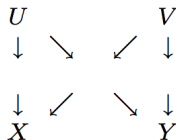
Path Counting: Half Sibs



- ▶ The common ancestor of half sibs X and Y is V . The path linking X, Y to their common ancestor is XVY and this has $n = 3$ individuals. Therefore

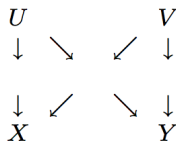
$$\theta_{XY} = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

Path Counting: Full Sibs



- ▶ Calculate the kinship coefficient for full sibs using path counting

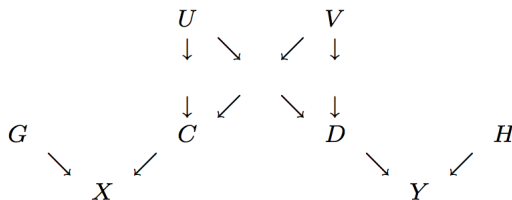
Path Counting: Full Sibs



- ▶ The common ancestors of full sibs X and Y are U and V . The paths linking X and Y to their common ancestors are XUY and XVY and these each have $n = 3$ individuals. Therefore

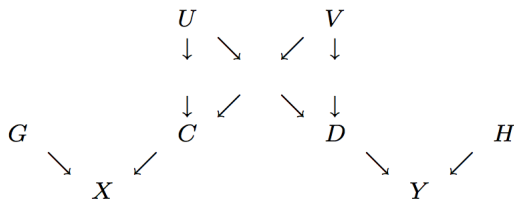
$$\theta_{XY} = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

Path Counting: First Cousins



- ▶ Calculate the kinship coefficient for first cousins using path counting

Path Counting: Full First Cousins



- ▶ The common ancestors of first cousins X and Y are U and V . The paths linking X and Y to their common ancestors are $XCUDY$ and $XCVDY$ and these each have $n = 5$ individuals. Therefore

$$\theta_{XY} = \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 = \frac{1}{16}$$

Coefficients of Relatedness

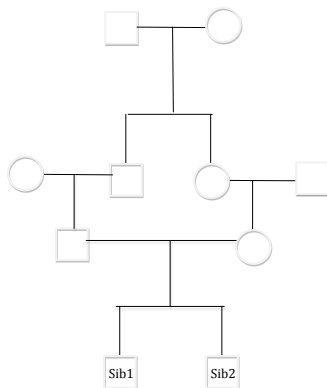
- ▶ Coefficients of Identity sharing probabilities and kinship coefficients for a few outbred relative pairs

IBD Sharing Probabilities for Outbreds

Relationship	Δ_7	Δ_8	Δ_9	θ
Parent-Offspring	0	1	0	$\frac{1}{4}$
Full Siblings	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
Half Siblings	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$
Uncle-Nephew	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$
First Cousins	0	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{1}{16}$
Double First Cousins	$\frac{1}{16}$	$\frac{6}{16}$	$\frac{9}{16}$	$\frac{1}{8}$
Second Cousins	0	$\frac{1}{16}$	$\frac{15}{16}$	$\frac{1}{64}$
Unrelated	0	0	1	0

- ▶ Note that for outbred populations $\Delta_7 + \Delta_8 + \Delta_9 = 1$
- ▶ Δ_7 , Δ_8 , and Δ_9 are often called IBD sharing probabilities

9 Coefficients of Identity: Pedigree Example



- ▶ The pedigree above contains two siblings from parents who are first cousins. Calculate the 9 condensed coefficients of identity for the siblings.

9 Coefficients of Identity: Pedigree Example

- ▶ For Δ_1 we must have that the parents of the sibs share 1 allele IBD and each parent passes down the IBD allele. The parents share 1 allele IBD with probability $\frac{1}{4}$, so

$$\Delta_1 = \left(\frac{1}{4}\right) \left(\frac{1}{2}\right)^4 = \left(\frac{1}{4}\right) \left(\frac{1}{16}\right) = \frac{1}{64}$$

- ▶ What is Δ_2 ? $\Delta_2 = 0$ since it is not possible for both siblings to be inbred at the locus but with the alleles in sibling 1 to not be IBD with the alleles of sibling 2 at the locus.
- ▶ Homework 3 will include the calculation of all 9 coefficients of identity as well as the kinship coefficients and inbreeding coefficients for the siblings in this pedigree.