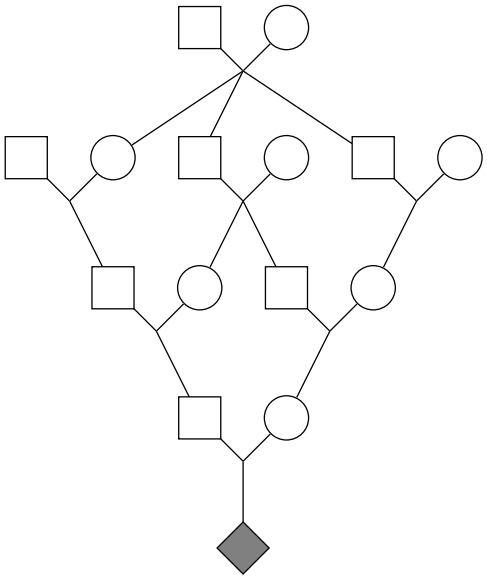
Chapter 2: Genes in Pedigrees

2.1.1 GRAPHICAL REPRESENTATION OF PEDIGREES:

- Three graphical representations.
  The parent-offspring links.
  —often animal pedigrees.
  —too many crossing lines.
  The sibship representation.
  —often human pedigrees
  —“tramlines” (J.H.Edwards)
  The marriage-node graph. — see figure (Cannings et al., 1978)
- Founders and non-founders (no half-founders): assumed unrelated.
- Gender: male, female, and unknown. (square, circle, diamond)
- Shading or labelling of individuals



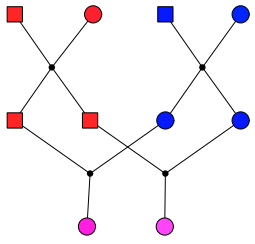
2.1.2 SPECIFICATION OF PEDIGREES:

- Unique individual identifiers ("names")
- Parent-offspring trios.
  (default: ind, dad, mom)
- Specification of founders.
  (parent "names" =0)
- Gender: male, female, and unknown. (1, 2, 0) or (M, F, U)
- Phenotypic, covariate, and marker data.
- "Chronological" (partial) ordering of pedigrees.

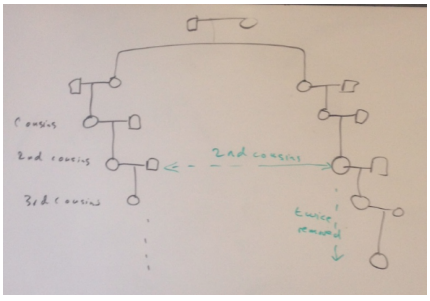
name	dad	mom	sex	other data
101	0	0	1	——-
102	0	0	2	——-
201	101	102	2	——-
204	101	102	1	——-
206	101	102	1	——-
fred	0	0	1	——-
203	0	0	2	
joe	fred	201	1	——-
jane	204	203	2	
dave	204	203	1	
hugh	joe	jane	1	
etc				

2.1.3 TYPES OF RELATIONSHIP:

- Related : individuals having a common ancestor (implies a biological relationship)
- Inbred: individuals whose parents are related (implies the maternal and paternal genes can descend from single ancestral gene).
- Unilateral (one-sided) and bilateral (two-sided) relationships:
unilateral: half-sibs, aunt, niece, cousins
bilateral: sibs, double first cousins, etc.
- Cousin-type relationships:
Half, full, and double cousins:

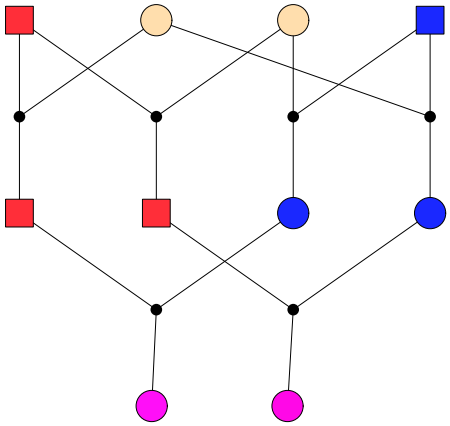


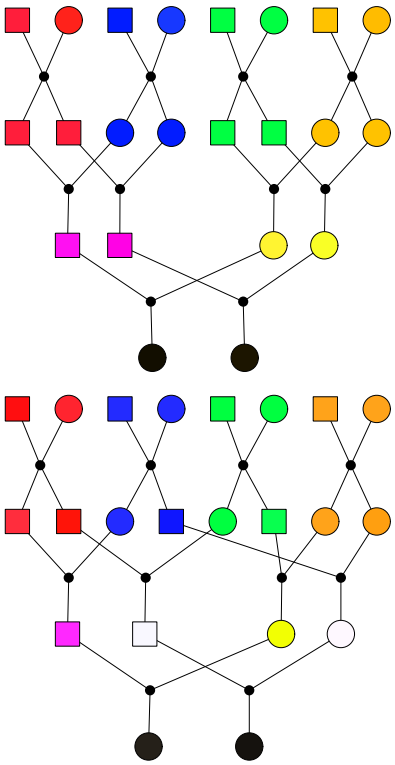
n^{th} cousins k times removed



2.1.4 More complex relationships:

Quadruple half first cousins;
and
quadruple second cousins





2.2.1 GENE IDENTITY BY DESCENT (*ibd*):

• RELATIVES ARE SIMILAR because they have *ibd* genes, that are copies of the same gene in a common ancestor.
NOTE: *ibd* is defined relative to given pedigree or time point

• **Basic assumption**:
— *ibd* genes are of the same allelic type; ignores mutation
— non-*ibd* genes are of independent types; ignores structure beyond the pedigree.
Recall the basic Mendelian genetics examples of Homework-1.

• **Framework for analysis of genetic data on pedigrees:**
A pedigree or relationship determines probabilities of *ibd*,
which determine probabilities of joint genotypes
which determine probabilities of joint phenotypes
that is, similarity among relatives.

2.2.2 KINSHIP and INBREEDING:

• The simplest pedigree-defined probabilities of gene *ibd* are the coefficients of kinship (ψ) and inbreeding (f), which measure *ibd* between two genes.

$$\begin{aligned} \psi(B,C) &= \Pr(homologous\ genes\ segregating \\ &\quad from\ B\ and\ C\ are\ ibd) \\ f(B) &= \Pr(homologous\ genes\ in\ B\ are\ ibd) \\ &= \psi(M_B, F_B) \end{aligned}$$

where M_B and F_B are the parents of B.

• Note at a single locus:

$$\text{randomly chosen from} \equiv \text{segregating from}$$

but for two of more loci we must define what "randomly chosen" should mean, and the relevant definition is that of Mendelian segregation.

2.2.3 KINSHIP EXAMPLES via PATH COUNTING:

Half sibs: $(1/2) \times (1/2) \times (1/2) = 1/8$

Two genes from an inbred (f) parent:
$1 \times f + (1/2) \times (1-f) = (1/2)(1+f)$

Half sibs with inbred (f) parent: $(1+f)/8$

Full sibs: $1/8 + 1/8 = 1/4$

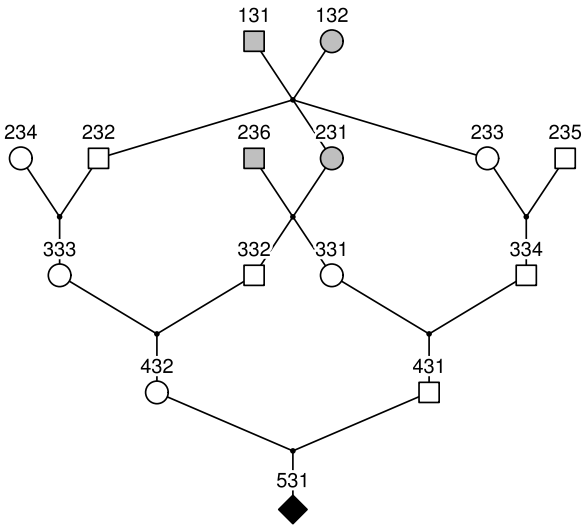
First cousins: $(1/4) \times (1/2) \times (1/2) = (1/16)$

Double first cousins: $1/16 + 1/16 = 1/8$

General formula (Wright, 1922):

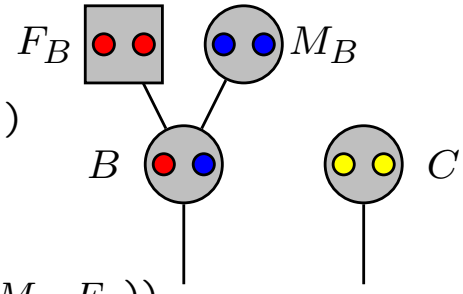
$$\psi = \sum_A \sum_{\mathcal{P}(A)} \frac{1}{2}(1+f_A) \left(\frac{1}{2}\right)^{n_1(\mathcal{P}(A))+n_2(\mathcal{P}(A))}$$

EXAMPLE: The JV pedigree (Goddard et al., 1996):



2 ancestors, each with 3 paths, each with $n_1 = n_2 = 3$:
and 2 ancestors, each with 1 path, each with $n_1 = n_2 = 2$.
$2 \times 3 \times (\frac{1}{2})^7 + 2 \times 1 \times (\frac{1}{2})^5 = 7/64$

2.2.4 RECURSIVE METHOD:



$$\psi(B,C) = \frac{1}{2}(\psi(M_B,C) + \psi(F_B,C))$$

provided B is not C nor an ancestor of C

$$\psi(B,B) = \frac{1}{2}(1 + f_B) = \frac{1}{2}(1 + \psi(M_B,F_B))$$

Boundary conditions:

$$\psi(A,A) = \frac{1}{2} \text{ and } \psi(A,C) = 0$$

if A is a founder, and not an ancestor of C

Expanding up the JV pedigree, among the grandparents, we have 3 first-cousin pairs and a sib pair. The kinship of first cousins is 1/16, and of sibs is 1/4, so overall we have

$$\frac{1}{4}(3\frac{1}{16} + \frac{1}{4}) = \frac{7}{64}$$

2.2.5 INBREEDING and GENOTYPE FREQUENCIES:

• MYTH: Recessive diseases are more frequent in genetic isolates. This is because isolates are "more inbred"

• TRUTH-1: the more inbred individuals within any population have higher probability of being affected.

• Consider a recessive disease allele a with freq q, and an individual with inbreeding coefficient f (aa individuals are affected.)

$$\begin{aligned}\Pr(aa) &= q^2(1-f)+qf &= q^2+fq(1-q)\\ \Pr(Aa) &= 2q(1-q)(1-f)\\ \Pr(AA) &= (1-q)^2+fq(1-q)\end{aligned}$$

See population mixtures and Wahlund variance (see 1.2.4),
For two alleles, see homework 2, # 4: $\sigma_f^2 = \sum_i \alpha_i(q_i-\bar{q})^2$:

$$\begin{aligned}\Pr(aa) &= q^2+\sigma_f^2 \text{ and } \Pr(AA) &= (1-q)^2+\sigma_f^2\\ \Pr(Aa) &= 2q(1-q)-2\sigma_f^2\end{aligned}$$

Population subdivision vs inbreeding:

• In population subdivision, people marry those more similar, hence more homozygosity in offspring.

• In inbreeding, people marry relatives, and hence more similar, and hence ... Inbreeding is a form of population subdivision.

Autozygosity (*ibd*) vs. inbreeding

• Autozygous ≡ having *ibd* genes
• inbred ≡ having non-zero prob of being autozygous

$$\Pr(\textit{ibd} \mid \text{affected}) = \frac{qf}{q^2+fq(1-q)} = \frac{f}{q+f(1-q)}$$

which $\geq f$, increases as q decreases, and $\to 1$ as $q \to 0$.

TRUTH-2: In a population with varying levels of inbreeding, the affected people have higher probability of being inbred.

Suppose a proportion α of the population (Pop_1) has inbreeding coefficient f and others (Pop_2) are not inbred:

$$\begin{aligned}
\Pr(\text{affected } aa) &= (1-\alpha)q^2 + \alpha(q^2 + fq(1-q)) \\
&= q^2 + \alpha f q(1-q) \\
\Pr(Pop_1 \mid \text{affected}) &= \frac{\alpha(q^2 + fq(1-q))}{q^2 + \alpha f q(1-q)} \\
&= \frac{\alpha(q + f - fq)}{(q + \alpha f - \alpha f q)}
\end{aligned}$$

which is always $\geq \alpha$ and $\to 1$ as $q \to 0$.

TRUTH-3: The affected inbred people in the population have higher probability of being autozygous (*ibd*):

$$\Pr(\textit{ibd} \mid \text{affected}) = \frac{\alpha f q}{q^2 + \alpha f q(1-q)} = \frac{\alpha f}{q + \alpha f(1-q)}$$

Same form as before with f now becoming αf.

Blank slide: