NONRANDOM MATING, GENETIC DRIFT, & MUTATION

NONRANDOM MATING/INBREEDING

READING: Nielsen & Slatkin, pp. 13–16, 59-63, 198-205

•Will distinguish two types of nonrandom mating:

(1) **Assortative mating:** mating between individuals with similar phenotypes or among individuals that occur in a particular location.

(2) **Inbreeding:** mating between related individuals.

– Both types of nonrandom mating may have similar consequences since individuals with similar phenotypes often have similar genotypes.

– It is often difficult to separate cause from effect.

• E.g., individuals with similar phenotypes may mate because
a) phenotypic assortative mating occurs;
b) mating with relatives is preferred;
c) matings are primarily based on proximity.

• Population subdivision: The Wahlund Effect

– It turns out that population subdivision *per se* can effect the distribution of genotypes in the entire population.

– Consider a locus with 2 alleles (*A* and *a*) and a collection of *isolated* subpopulations numbered 1, 2, 3, ...

– Let the frequencies of *A* and *a* in subpopulation *i* be p_i and q_i.

– Assuming random mating within each (isolated) subpopulation:
• Freq(*AA*) in subpopulation *i* = p_i^2
• Freq(*Aa*) in subpopulation *i* = $2p_iq_i$
• Freq(*aa*) in subpopulation *i* = q_i^2

– Let $\bar{p} = \underset{i}{\text{Avg}}(p_i)$ = average freq. of *A* across all subpopulations.

• Likewise, let $\bar{q} = \underset{i}{\text{Avg}}(q_i)$ = 1 - $\bar{p}$

– What is the average frequency of each genotype over all the subpopulations?

• Consider *AA* homozygotes first:

– From Fun Facts: $\mathrm{Var}(X) = E(X^2) - [E(X)]^2$ so $E(X^2) = [E(X)]^2 + \mathrm{Var}(X)$

– So, $\underset{i}{\mathrm{Avg}}(p_i^2) = \overline{p^2} = \bar{p}^2 + \mathrm{Var}(p)$

• Similarly, for *aa* homozygotes:

– $\underset{i}{\mathrm{Avg}}(q_i^2) = \overline{q^2} = \bar{q}^2 + \mathrm{Var}(q) = \bar{q}^2 + \mathrm{Var}(p)$ since Var(q) = Var(1 – p) = Var(p).

• Finally, for *Aa* heterozygotes:

– $\underset{i}{\mathrm{Avg}}(2p_iq_i) = 1 - \overline{p^2} - \overline{q^2} = 1 - \bar{p}^2 - \bar{q}^2 - 2\mathrm{Var}(p) = 2\bar{p}\bar{q} - 2\mathrm{Var}(p)$.

– A Thought Experiment:

• Suppose genotypes were randomly sampled from a population whose substructure was unknown.

– The frequencies of *A* and *a* in sample would be $\bar{p}$ and $\bar{q}$.

– With random mating, would then *expect* to find genotypes in proportions $AA : Aa : aa = \bar{p}^2 : 2pq : \bar{q}^2$.

– But, the genotype frequencies *observed* would be $AA : Aa : aa = \overline{p^2} : \overline{2pq} : \overline{q^2} = \bar{p}^2 + \mathrm{Var}(p) : 2\bar{p}\bar{q} - 2\mathrm{Var}(p) : \bar{q}^2 + \mathrm{Var}(p)$.

• I.e., would find an excess of homozygotes and a deficit of heterozygotes, compared to expectations.

– Why? Simply because of population subdivision and, in particular, variance in allele frequencies across subpopulations.

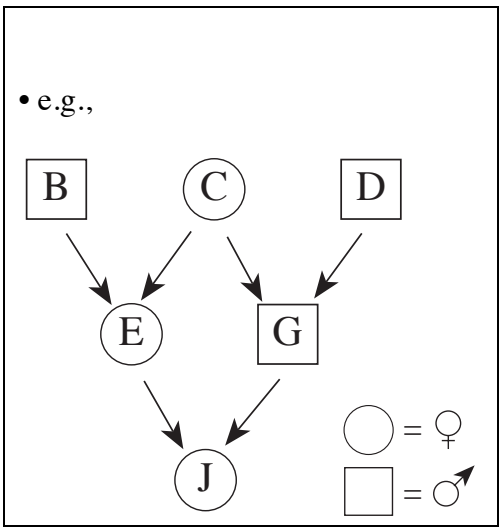
– Given across-subpopulations differences in allele frequencies, the apparent excess in homozygotes and deficit of heterozygotes from what is *expected* were the entire population to mate at random defines what is called the **Wahlund Effect**.

– The Wahlund effect is a common "cause" of non-conformity to Hardy-Weinberg expectations in population samples.

• **INBREEDING**

– Will now consider the genetic consequences of mating between relatives: "inbreeding."

• e.g.,



• In 1920's, Sewall Wright invented an ingenious approach to tracking genotypes through pedigrees based on the probability that allele copies are "identical by descent."

• Will follow the French geneticist Malecot's re-working of Wright's method here.

– **Identity by descent (IBD):** Two alleles are "identical by descent" if

(1) both are descended from the same allele in a common ancestor
or
(2) one allele is descended from the other.

• Will mean IBD relative to a specific base population (whose alleles are *deemed* to be not IBD).

• Definition: The **inbreeding coefficient**, f_J, of an individual *J* is the probability that its two gene copies at a locus are identical by descent.

• Once f_J is known, it's not hard to find the probabilities that *J* is *AA*, *Aa*, or *aa*:

– Consider a randomly chosen individual:

• With probability f_J, both gene copies in that individual are IBD.
– Then *both* will be *A* if the allele they were copied from in the base population were *A*.
• But, *A* occurs with frequency *p* in the base population, so the probability of being an *AA* given both genes are IBD is *p*.

$\Rightarrow$ the probability of getting two *A* alleles that are IBD is $f_J p$.

– Likewise, the probability of being *aa* with both genes are IBD is $f_J(1-p)$.

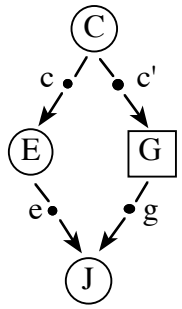
• With probability $1 - f_J$, the two genes in an individual will *not* be IBD.
– Must have descended from different allele copies in the base population.

– Assuming the copies are made independently, then with probability $p \times p = p^2$, the copied alleles are both *A* (genotype *AA*), etc.

• Putting this all together, have:

$$P_{AA} = (1 - f_J)p^2 + f_J p$$
$$P_{Aa} = (1 - f_J)2p(1 - p)$$
$$P_{aa} = (1 - f_J)(1 - p)^2 + f_J(1 - p)$$

• Great! So now only need to determine f_J. Thanks to Wright, this is very easy to do.



– E.g., Let's find f_J in the pedigree above.

– Need only concentrate on the central part of the pedigree:

$$f_J = \text{Prob}(e = c) \times \text{Prob}(c = c') \times \text{Prob}(c' = g)$$
$$= 1/2 \times 1/2 \times 1/2$$
$$= 1/8$$

– So the expected genotype frequencies for this pedigree are:

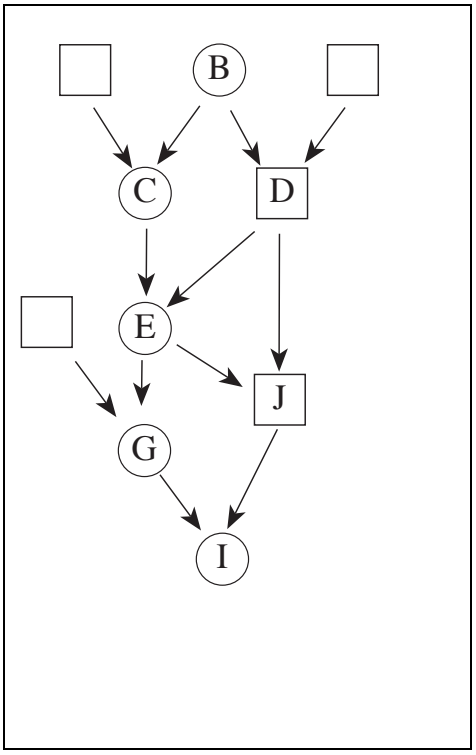
$$P_{AA} = (7/8)p^2 + 1/8\,p$$
$$P_{Aa} = (7/8)2p(1 - p)$$
$$P_{aa} = (7/8)(1 - p)^2 + 1/8(1 - p)$$

– ASIDE: What is the average frequency of *A* among individuals with inbreeding coefficient f_J?

$$\text{Freq}(A) = P_{AA} + \tfrac{1}{2}P_{Aa} = \left[(1 - f_J)p^2 + f_J p\right] + \tfrac{1}{2}\left[(1 - f_J)2p(1 - p)\right]$$
$$= (1 - f_J)\left[p^2 + p - p^2\right] + f_J p = (1 - f_J)p + f_J p$$
$$= p$$

• Inbreeding does not affect allele frequencies *on average*, but does affect the probabilities that 2 *A*'s or 2 *a*'s co-occur in an individual.

– **Computing Inbreeding Coefficients** (in general)



• Say we want to find f_J in the pedigree at right:

• Rules:

(1) Enumerate each loop

(2) Each loop must

a) go through each individual no more than once

b) only change from up to down once

(3) Multiply by ½ for each passage through an individual

– If the passage through an individual involves a change of direction (up/down) multiply by $(1+f)/2$ instead of 1/2, where f is the inbreeding coefficient for that individual.

(4) Add the probabilities of each loop.

• For the above example:

$$\begin{aligned} f_J &= \tfrac{1}{2}^{C}\cdot\tfrac{1}{2}^{E}\cdot\tfrac{1}{2}^{C}\cdot(1+0^{B})/2\cdot\tfrac{1}{2}^{D}\cdot\tfrac{1}{2}^{J} \\ &+ \tfrac{1}{2}^{C}\cdot\tfrac{1}{2}^{E}\cdot(1+0^{D})/2\cdot\tfrac{1}{2}^{J} \\ &+ \tfrac{1}{2}^{C}\cdot(1+f_E^{E})/2\cdot\tfrac{1}{2}^{J} \\ &= 1/64+1/16+(1+1/8)/8 = 1/64+1/16+9/64 \\ &= 7/32 \end{aligned}$$

• Some forbidden loops: IGECBDEGI (goes through G twice)
IJDBCEGI (already counted)
IGECBDEJI (loop ECBDE already accounted for in f_E)

– Evolutionary Application: **Kin Selection**

• Probability that *two* individuals share an allele descended from a common ancestor is called the **kinship coefficient** or **coefficient of consanguinity**.

• Kinship coefficient between individuals A and B is denoted F_{AB}.

• What is F_{GJ} in the last example?

– Clearly, it is the inbreeding coefficient of their offspring I, $f_I = 7/32$.

• The connection between f and F_{AB}:

– The kinship coefficient of two individuals is equal to the inbreeding coefficient of their (perhaps hypothetical) offspring
• e.g., $F_{\text{mother,daughter}} = 1/4$ (assuming unrelated, non-inbred parents).

• Kinship coefficients useful in studying evolution of social (especially "altruistic") traits.

– In 1964, W. D. Hamilton proposed a rule-of-thumb to determine whether a rare allele will be favored by selection:

• A rare mutation which affects the fitness of its carrier and others will spread if $br > c$ where
b = "benefit" = increase in fitness to recipients of action
c = "cost" = loss in fitness to actor
r = "degree of relatedness".
– can use above rules to find r: $r = 2\,F_{AB}$

– Another view of f_J:

• f_J is the proportion by which heterozygosity is reduced relative to a random mating group with the same allele frequencies.
– If $f_J = 0$, the population is in Hardy-Weinberg proportions.
– If $f_J = 1$, all individuals are homozygous.
– In general: $\text{Het}_{f_J} = (1 - f_J)\text{Het}_0$ where $\text{Het}_0 = 2p(1-p) = 2pq$

• Consider, e.g., the Wahlund effect:
– Overall frequencies of A and a are $\bar{p}$ and $\bar{q}$.
– If individuals from all subpopulations mate at random, expect to find $2\bar{p}\,\bar{q}$ heterozygotes.
– Would actually observe $2\bar{p}\bar{q} - 2\text{Var}(p)$ heterozygotes.
– Thus, in this case $2\bar{p}\bar{q} - 2\text{Var}(p) = (1-f)2\bar{p}\bar{q}$.
– In order for this to be true, $f = \text{Var}(p)/\bar{p}\bar{q}$. [Suggested exercise: show this.]
– Note: $f \neq 0$ even though there is no deficit of heterozygotes *within* subpopulations and no obvious "inbreeding"!