BIOL 502 Population Genetics Spring 2017

Week 8 Inbreeding

Arun Sethuraman

California State University San Marcos

Table of contents

Inbreeding Coefficient

Inbreeding Coefficient

Inbreeding

Non-random mating between individuals that are related by recent descent. E.x. self fertilization in plants, first-cousin mating in humans, brother-sister mating in Drosophila.

Inbreeding Coefficient, F

The probability that two alleles at a locus in an inbred individual are identical by descent (IBD), or alternately, the overall proportion of genes in an inbred individual that have alleles that are IBD.

- Consider an arbitrary time in the past, generation 0, where all alleles are *distinct*, i.e. *not* IBD.
- In this generation, let the genotype of an individual at a locus be $\alpha_1 \alpha_2$, such that α_1 and α_2 are not IBD.
- In generation 1, an individual has any one of 4 possible genotypes with the probabilities $\frac{1}{4}\alpha_1\alpha_1, \frac{1}{4}\alpha_1\alpha_1, \frac{2}{4}\alpha_1\alpha_2, \frac{1}{4}\alpha_2\alpha_2$.
- Therefore the inbreeding coefficient in a *selfing* system is $F = \frac{1}{2}$.

Genotype Frequencies

- Consider a biallelic genetic locus, A, such that it has two alleles A and a at frequencies p and q respectively in a population, such that $p + q = 1$.
- Let F be the probability of IBD (or autozygosity), or in other words, the reduction in heterozygosity due to inbreeding (or increase in homozygosity due to inbreeding).
- If you randomly sample two alleles at a locus, they can either be autozygous (with a probability F), or allozygous (with a probability $1 - F$).
- Genotype frequencies after one generation of inbreeding can then be written as $P(AA) = p^2(1 - F) + pF$, $P(Aa) = 2pq(1 - F)$, $P(aa) = q^2(1 - F) + qF$.
- If $F = 0$, the population is in HWE, and randomly mating.
- If $F = 1$, the population is completely inbred.
- How about allele frequencies?
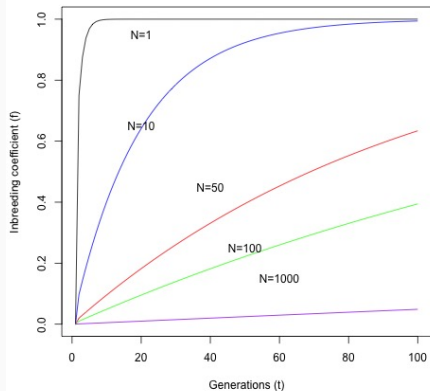
Reduction in heterozygosity

- Let H_I be the probability that a gene in an inbred individual I is heterozygous, and H_S be the proportion of heterozygous genotypes expected under random mating in the *subpopulation* of which I is a member.
- $H_I = 2pq - 2pqF$, and $H_S = 2pq$
- $F_{IS} = \frac{H_S - H_I}{H_S} = \frac{2pq - (2pq - 2pqF)}{2pq} = F$

Inbreeding in finite populations

- Now if the population is of finite size, $2N$, the probability that any two alleles are IBD in the first generation is $F = \frac{1}{2N}$.

- $f_{t+1} = \frac{1}{2N} + (1 - \frac{1}{2N})f_t$

- In the first generation, $f_0 = 0$, and hence the recursion can be written as $f_t = 1 - (1 - \frac{1}{2N})^t$

- Expected heterozygosity will decrease according to $H_t = H_0(1 - \frac{1}{2N})^t$.

Inbreeding coefficient variation with N_e

```
#jpeg("fvtimevaryingN.jpeg")
N=1
ft[1]=0
for (i in 2:100){ft[i]=1-(1-1/(2*N))^i}
plot(ft,type="l",xlab="Generations (t)",
ylab="Inbreeding coefficient (f)")
N=10
for (i in 2:100){ft[i]=1-(1-1/(2*N))^i}
points(ft,type="l",col="blue")
N=50
for (i in 2:100){ft[i]=1-(1-1/(2*N))^i}
points(ft,type="l",col="red")
N=100
for (i in 2:100){ft[i]=1-(1-1/(2*N))^i}
points(ft,type="l",col="green")
N=1000
for (i in 2:100){ft[i]=1-(1-1/(2*N))^i}
points(ft,type="l",col="purple")
text(20,0.95,"N=1")
text(20,0.65,"N=10")
text(40,0.45,"N=50")
text(50,0.25,"N=100")
text(60,0.15,"N=1000")
dev.off()
```

Mating Systems

Complete Self Fertilization

- Recall that in a selfing system, $F = \frac{1}{2}$ after one generation.
- In other words, the reduction in heterozygosity after one generation of selfing is by a factor of $\frac{1}{2}$.
- $H_{t+1} = \frac{1}{2} H_t$
- In other words, heterozygosity declines rapidly to 0.
- $f_{t+1} = \frac{1}{2}(1 + f_t)$

Mixed Mating

- Consider a system with mixed selfing and outcrossing.
- If S is the proportion of the population that is selfing, $1 - S$ is the proportion that undergoes outcrossing.
- $f_{t+1} = S(\frac{1}{2})(1 + f_t) + (1 - S) \times 0 = \frac{1}{2}S(1 + f_t)$
- At equilibrium, setting $f_{t+1} = f_t$, we get $\hat{f} = \frac{S}{2-S}$
- Similarly, $H_{t+1} = (1 - S)H_0 + S(\frac{1}{2}H_t)$
- At equilibrium, $H_{t+1} = H_t$, and $\hat{H} = 2(\frac{1-S}{2-S})H_0$.

Mixed Mating - Heterozygosity decline

```
jpeg("hetselfing.jpeg")
H<-c(1:20)
H[1]=1.0
for (i in 2:20){S=0.1;
H[i]=(1-S)*H[1]+S*0.5*H[i-1]}
plot(H,xlab="Generations",ylab="Heterozygosity",
ylim=c(0,1),type="l")
for (i in 2:20){S=0.25;
H[i]=(1-S)*H[1]+S*0.5*H[i-1]}
points(H,type="l")
for (i in 2:20){S=0.5;
H[i]=(1-S)*H[1]+S*0.5*H[i-1]}
points(H,type="l")
for (i in 2:20){S=0.75;
H[i]=(1-S)*H[1]+S*0.5*H[i-1]}
points(H,type="l")
for (i in 2:20){S=0.9;
H[i]=(1-S)*H[1]+S*0.5*H[i-1]}
points(H,type="l")
for (i in 2:20){S=1.0;
H[i]=(1-S)*H[1]+S*0.5*H[i-1]}
points(H,type="l")
text(10,1.0,"s=0.1")
text(10,0.85,"s=0.25")
text(10,0.65,"s=0.5")
text(10,0.45,"s=0.75")
text(10,0.25,"s=0.9")
text(10,0.05,"s=1.0")
dev.off()
```
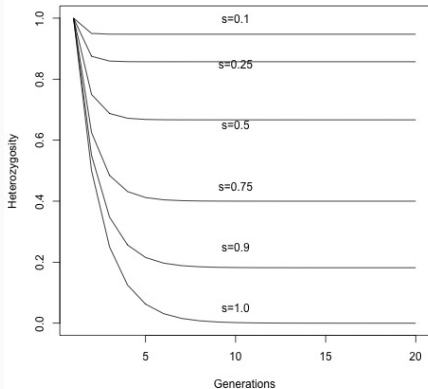
Consanguinity and Inbreeding

Coefficient of Consanguinity

Consanguinity

Two individuals are said to be consanguineous if they have a common ancestor.

Coefficient of Consanguinity

Probability that two alleles, chosen at random from each of two individuals are identical by descent (IBD).

Pick an allele at random from individual, X. Pick an allele at random from the same locus in individual Y; the probability that they are IBD is the coefficient of consanguinity/kinship/coancestry.

$g_{XY} = \sum_{allalleles} \frac{1}{2} Pr(X = A_i) \times \frac{1}{2} Pr(Y = A_i)$, where A_i is the ith allele that is possibly IBD from the common ancestor, $Pr(X = A_i)$ and $Pr(Y = A_i)$ are the probability that individuals X or Y has allele A_i.

Full siblings

For example, if two individuals are full siblings with common parents A and B, $g_{XY} = \sum_{allalleles} \frac{1}{2} (\frac{1}{2}) \times \frac{1}{2} (\frac{1}{2}) = \frac{1}{4}$

Coefficient of Relatedness

Expected proportion of alleles that are IBD in two individuals.

1. Draw up the pedigree as a tree (showing only a single parent, and children).
2. Mark the shortest possible path between two family members and count the number of steps between them. Each step across generations (parent to child) or between siblings counts as one.
3. Once you have this number of steps (n), $r_{XY} = \frac{1}{2}^n$, or $\frac{1}{2}$ multiplied by itself n times.

Full siblings

If two individuals X and Y are full siblings, with common parents A and B, there's only one path - $M - N$, so $r_{XY} = \frac{1}{2}^1 = \frac{1}{2}$

Inbreeding Coefficient from Pedigrees

1. List all common ancestors
2. List all pathways through each common ancestor - start with a parent, work back to the common ancestor, then forward to the other parent.
3. One individual can only be present once in a pathway, but can be present in multiple pathways, let number of individuals in the ith pathway be n_i.
4. If the common ancestor is not inbred, the partial inbreeding coefficient of the ith pathway is $f_i = (\frac{1}{2})^{n_i}$
5. If the common ancestor is itself inbred, multiply the partial inbreeding coefficient by one plus the inbreeding coefficient of the common ancestor: $f_i = (\frac{1}{2})^{n_i}(1 + f_{ai})$
6. Repeat steps 4 and 5 for all possible pathways.
7. Sum the partial inbreeding coefficients through all possible pathways to get $f = \sum_{i=1}^{p}(\frac{1}{2})^{n_i}(1 + f_{ai})$

Inbreeding Depression

Inbreeding Depression

Inbreeding Depression

The detrimental effects of inbreeding, due to accumulation of deleterious mutations.

- Let the fitnesses of the genotypes AA, Aa, aa be 1, $1 - hs$ and $1 - s$ respectively.
- Assume that a is a deleterious allele, that is partially recessive, and that the heterozygote has intermediate fitness ($0 < h < 1$).
- Then mean fitness under random mating is
  $\bar{w}_O = p^2 + 2pq(1 - hs) + q^2(1 - s)$
- If the population is inbreeding, then we get
  $\bar{w}_I = (p^2 + pqf) + 2pq(1 - f)(1 - hs) + (q^2 + pqf)(1 - s)$.
- Rewriting this, $\bar{w}_I = \bar{w}_O + pqf[1 - 2(1 - hs) + (1 - s)]$.
- Inbreeding depression occurs when $\bar{w}_I < \bar{w}_O$, i.e. if
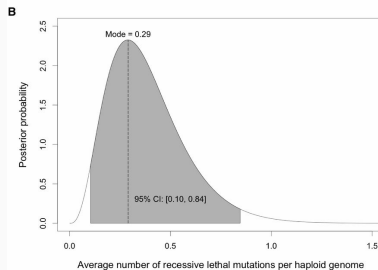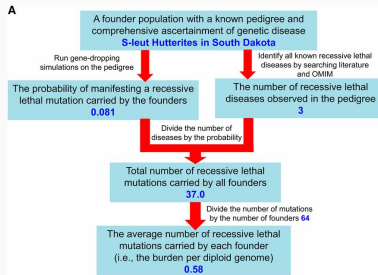 $1 - 2(1 - hs) + (1 - s) < 0$, or $h < 0.5$.

The Inbreeding of Charles II of Spain
(1661-1700)
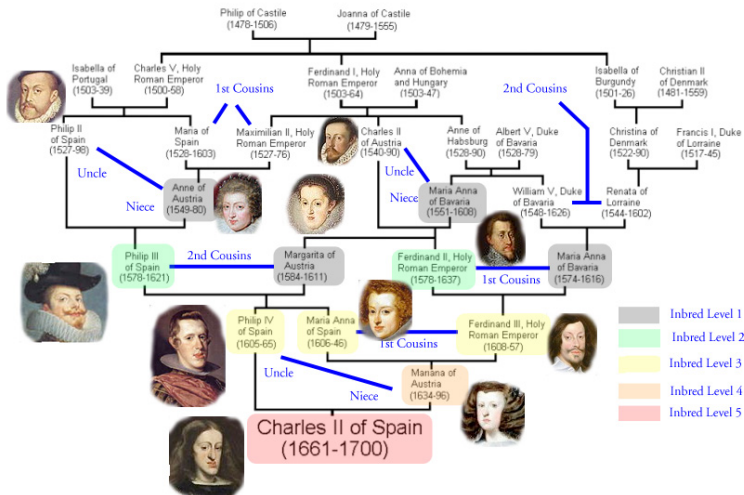
15

Methods to detect Inbreeding

Summary Statistics

F statistics

Can be computed using the `Fst` function in the `pegas` package.

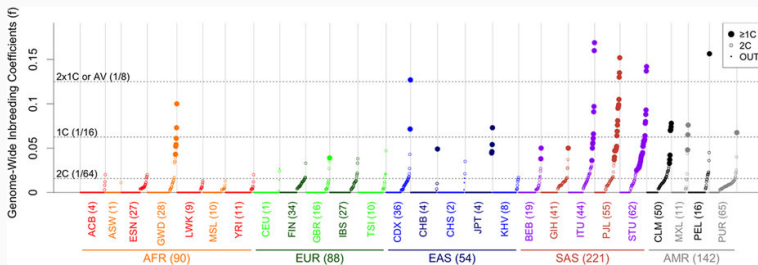**Heterozygosity/homozygosity**

Can be computed using the `heterozygosity` function in `pegas` package, then homozygosity = 1-heterozygosity.

Relatedness

Can be computed using the `related` package in R, see DOI: 10.1111/1755-0998.12323 (Pew et al. 2015).

Runs of Homozygosity

- Longer "runs" expected under recent inbreeding
- Why?
- How about short runs?
- Commonly used algorithms - PLINK, BEAGLE

Figure 3