

GENE613 - Exam 1

Daniel Osorio - dcosorih@tamu.edu

Department of Veterinary Integrative Biosciences

Texas A&M University

1. *Define equilibrium allele frequency:* Is a principle proposed by Hardy and Weinberg in 1908, stating that the allele frequency in a population will remain constant from one generation to the next in the absence of selection, mutation or migration.
2. *In the context of single locus with 2 alleles initially with genotypes in Hardy-Weinberg proportions $f(A_1) = p$; $f(A_2) = q$. A proportion s of the A_2A_2 genotypes do not survive past early childhood. Show algebraically that after that proportion is gone (and before the next generation is produced) that the change in A_2 allele frequency is $\Delta q = \frac{psq^2}{1-sq^2}$*

$$\begin{aligned}\frac{psq^2}{1-sq^2} &= q - \frac{(1-sq)q}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{q(1-sq^2)}{1-sq^2} - \frac{(1-sq)q}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{q(1-sq^2) - (1-sq)q}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{q(1-sq^2) - (q - sq^2)}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{q(1-sq^2) - q + sq^2}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{(1-p)(1-sq^2) - q + sq^2}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{\cancel{1} - sq^2 \cancel{p} + psq^2 \cancel{q} + sq^2}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{\cancel{1} - \cancel{sq^2} + psq^2 + \cancel{sq^2}}{1-sq^2} \\ \frac{psq^2}{1-sq^2} &= \frac{psq^2}{1-sq^2}\end{aligned}\tag{1}$$

3. *In the context of a progeny test of an individual that may be heterozygous at a certain locus with dominant gene action, we assume that test mates and meiosis are independent. Explain these assumptions, including why they are necessary and give examples of both that would not be independent.*
 - a) *Test mates independence assumption:* This assumption implies that each mate will produce gametes in function of their genotype and those gametes do not are dependent of the gametes of other mate. This allows us to multiply the probabilities of each mate and their offspring in the used formula.

- b) *Meiosis independence assumption*: This assumption implies that all individuals will produce two types of gametes and each type will carry one of the alleles that conforms their genotype, this ensures that the allele proportion in the whole gametic output is the same as in the parentals. This allows us to compute and multiply the pre-test probabilities of nondetection.
- c) *Examples*: An exception to the test mates independence assumption is selection, a process where exist differential survival and reproduction of individuals due to differences in phenotype. An exception to the meiosis independence assumption for alleles that are located very close to one another on the same chromosome because of genetic linkage.
4. *In a herd of cattle, the homozygous recessive genotype at a single locus was responsible for horns (pp) (actually, inheritance of horns is a bit more complex than this, but we will assume this simplicity for now) and the frequency of the P allele was 0.4. Assuming HWE, if you decide to cull 40 % of the horned animals in the current calf crop*

- *What will be the change in $f(p)$ after one generation?*

```
> selection <- function(p, s) {
+   q <- 1 - p
+   return(((1 - s) * q ^ 2 + 0.5 * (2 * p * q)) / (1 - (s * q ^ 2)))
+ }
> selection(p = 0.4, s = 0) - selection(p = 0.4, s = 0.4)
[1] 0.06728972
```

- *What will be the change in $f(p)$ after two generations of culling all horned animals?*

```
> selection(p = 0.4, s = 0) - selection(p = (1 - selection(p = 0.4,
+                                                         s = 1)),
+                                     s = 1)
[1] 0.3272727
```

In mammalian species, there exist a recessive condition known as albinism, where the homozygous recessive genotype results in lack of pigmentation in skin, hair and eyes. The heterozygote has normal pigmentation, as does the homozygous dominant genotype. A male dog was produced in a litter from two normal appearing parents, but one of the five puppies in this litter was albino.

5. *What is the probability that the male puppy is homozygous dominant for this locus based on the above pedigree information?*

```
> punnett <- outer(c(A=0.5,a=0.5),c(A=0.5,a=0.5), "*")
> punnett["A", "A"]
[1] 0.25
```

6. *This same male dog was bred to 4 female dogs. Each of the females was itself a daughter of a documented heterozygote parent (how about the other parent? What you should assume about those?). Two litters with 2 and 3 normal appearance puppies, respectively, were produced. Is this a good test? Is this dog a carrier?*

- a) *How about the other parent? What you should assume about those?* As there is no information suggesting that the females are albino and one of the parents is homozygous, I assume that all the female mates are heterozygotes.

b) *Is this a good test? Is this dog a carrier?*

```
> litters <- c(2,3)
> PAa <- (punnett["A","a"] + punnett["a","A"])/(1-(punnett["a","a"]))
> PAA <- punnett["A","A"]/(1-(punnett["a","a"]))
> PDN <- 1-prod((3/4)^litters)
> PDN
```

```
[1] 0.7626953
```

```
> (PAA)/(((1-PDN)*PAa)+PAA)
```

```
[1] 0.6781457
```

No, it is not a good test, it only can detect a homozygote recessive in the 76 % of the cases. Exist the 67 % of probability that the dog be a carrier.

7. *What would be different about the above scenario of there was a single litter of 5 normal puppies produced from the same type of female parent, and why?*

```
> litters <- c(5)
> 1-prod((3/4)^litters)
```

```
[1] 0.7626953
```

No, there is not difference due to I assume that all the female mates are homozygotes, all will have the same genotype and the same *PND*.

Below is information from cotton breeding data set. A single locus has been identified with two alleles segregating (+,b) that impact lint yield (measured in lb per acre). $f(+) = p$; $f(b) = q$. Lint yields (lb per acre) for the genotypes were 825, 815, 785 for ++, +b, bb respectively.

8. *The following R script predicts population mean lint yields , breeding values, dominance deviations, breeding value variance, total genetic variance, and the ration of the breeding value to the total genetic variances under 3 different $f(b) = 0,2, 0,5, 0,8$. Correct any errors and interpret results. Discuss how your answer would change if $d = 0$ (i.e. yield of +b = 805)*

```
> q<-c(.2,.5,.8)
> phn<-c(825,815,785)
> a<-phn[1]-mean(c(phn[1],phn[3]))
> d<- phn[2]-mean(c(phn[1],phn[3]))
> answ<-matrix(0,13,3)
> rownames(answ)<-c("f(++)", "f(+b)", "f(bb)",
+                  "Mean", "BV ++", "BV +b", "BV bb",
+                  "DD ++", "DD +b", "DD bb", "V(A)", "V(G)", "V(A)/V(G)")
> colnames(answ)<-c("q=0.2", "q=0.5", "q=0.8")
> for (i in 1:3){
+   f.w<-(1-q[i])^2
+   f.wb<-2*(1-q[i])*q[i]
+   f.bb<-(q[i])^2
+   p <- (1-q[i])
```

```

+      M <- ((a*(p-q[i])) + 2*p*q[i])
+      alfa<- a+(d*(q[i]-p))
+      bv.wv<- 2*q[i]*alfa
+      bv.wb<- (q[i]-(1-q[i]))*alfa
+      bv.bb<- -2*(1-q[i])*alfa
+      dd.wv<- -2*(q[i]^2)*d
+      dd.wb<- 2*(1-q[i])*q[i]*d
+      dd.bb<- -2*((1-q[i])^2)*d
+      sigma2bv<- 2*p*q[i]*(alfa^2)
+      sigma2tg<-sigma2bv+(2*(1-q[i])*q[i])^2
+      ratio<-sigma2bv/sigma2tg
+      answ[,i]<-c(f.wv,f.wb,f.bb,
+                  M,bv.wv,bv.wb,bv.bb,
+                  dd.wv,dd.wb,dd.bb,
+                  sigma2bv,sigma2tg,ratio)
+ }
> print(answ)

```

	q=0.2	q=0.5	q=0.8
f(++)	0.64000	0.2500000	0.0400000
f(+b)	0.32000	0.5000000	0.3200000
f(bb)	0.04000	0.2500000	0.6400000
Mean	12.32000	0.5000000	-11.6800000
BV ++	5.60000	20.0000000	41.6000000
BV +b	-8.40000	0.0000000	15.6000000
BV bb	-22.40000	-20.0000000	-10.4000000
DD ++	-0.80000	-5.0000000	-12.8000000
DD +b	3.20000	5.0000000	3.2000000
DD bb	-12.80000	-5.0000000	-0.8000000
V(A)	62.72000	200.0000000	216.3200000
V(G)	62.82240	200.2500000	216.4224000
V(A)/V(G)	0.99837	0.9987516	0.9995269

The results show several values associated with the population in which the genotypic values were measured. As they are population measurements, they change in response to the q proportion changes. As the d value is not really far to 0 in comparison with a the most evident response is due to an additive activity of the alleles. If the d value is changed to 0 then the dominance values will turn to 0 and the proportion $\frac{V(A)}{V(G)}$ will be really close to 1.

9. Given that $P(Aa|CT) = \frac{[1-P(DN)]P(Aa)}{[1-P(DN)]P(Aa)+P(AA)}$ show that $P(AA|CT) = \frac{P(AA)}{[1-P(DN)]P(Aa)+P(AA)}$
10. Single locus with two alleles in mice. Record of dams.

	A_1A_1	A_1A_2	A_2A_2
Number of female mice evaluated:	110	240	131
Number of pups per litter:	6	7	8
Number of litters:	8	7	6
Total of pups:	48	49	48

```

> genotypicValues <- c(6, 7, 8)
> genotypicValues <- genotypicValues - mean(genotypicValues[c(1, 3)])

```

```

> a <- genotypicValues[1]
> d <- genotypicValues[2]
> A1A1 <- 110
> A1A2 <- 240
> A2A2 <- 131
> n <- sum(A1A1 + A1A2 + A2A2)
> A1 <- ((A1A1 + 0.5 * A1A2) / n)
> A2 <- (1-A1)
> c(
+   a = a,
+   d = d,
+   A1 = A1,
+   A2 = A2,
+   A1A1 = A1A1 / n,
+   A1A2 = A1A2 / n,
+   A2A2 = A2A2 / n,
+   MEAN = a * (A1 - A2) + 2 * A1 * A2
+ )

```

	a	d	A1	A2	A1A1	A1A2	A2A2
	-1.0000000	0.0000000	0.4781705	0.5218295	0.2286902	0.4989605	0.2723493
MEAN							
	0.5427060						

What is the genetic action at this locus? Give the allele and the genotype frequencies, and the mean of the population. Taking into account that $d = 0$ then, the genetic action at this locus is additive.