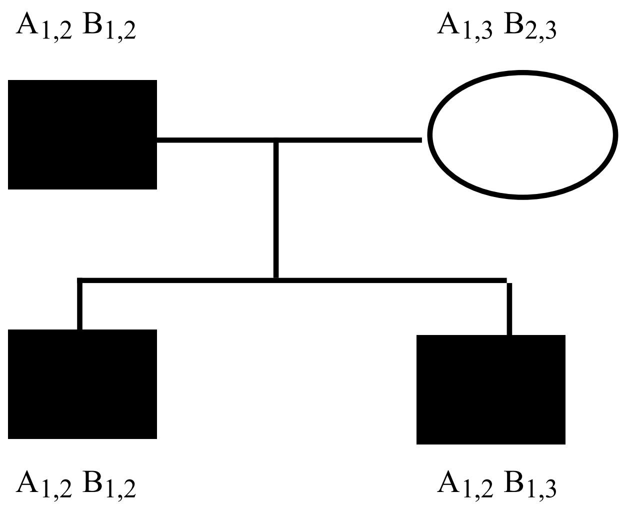
Molecular Genetics BSCI410 Homework 4; Lectures 13-17

Answer key

----

1. (2 points) Examine the following pedigree.



The A_1 alleles in the two brothers are identical by state (this just means that they are both A_1). The same is true of B_1. In either case can you infer that they are identical by descent? In other words, which alleles are identical by descent?

d) You can conclude that both A_1 and B_1 are identical by descent.

Both brothers inherited A_1 from their father and both brothers inherited B_1 from their father. The case of B_1 is pretty obvious. In the case of A_1, you know that they inherited the paternal A_1 allele because they did not inherit the paternal A_2 allele. Most of you got this correct.

(Questions 2-6): Consider two populations, 1 and 2, that differ at two unlinked loci, A and B. In each population a specific allele is fixed at each locus (*i.e.* all individuals in population 1 are homozygous for A_1 and B_1 -- they have the genotype $A_{1,1}$ $B_{1,1}$ -- while all individuals in population 2 have the genotype $A_{2,2}$ $B_{2,2}$).

First, you cross a single male from population 1 with a single female in population 2. Of course, all of the F1 progeny are heterozygous at both loci, $A_{1,2}$ $B_{1,2}$.

2. (1 point) Is locus A at Hardy-Weinberg equilibrium in the F1 generation (your answer would be the same for locus B)?

No, there is a deficit of homozygotes. All of the progeny will be A_1/A_2

Locus A is not at Hardy-Weinberg equilibrium in the F1 generation; for Hardy-Weinberg equilibrium, $p^2 + 2pq + q^2 = 1$ and $p + q = 1$, but this is not the case here since there are no A_1A_1 or A_2A_2 homozygotes.

3. (1 point) What is the expected frequency of each of the four possible gametes transmitted from the F1 to the F2. The possible genotypes are A_1 B_1, A_1 B_2 , A_2 B_1 and A_2 B_2

0.25 for each genotype.

This follows from the facts that all of the F1 are $A_{1,2}$ $B_{1,2}$ and A and B are unlinked loci.

4. (1 point) Do the two alleles A_1 and B_1 show genetic association in this F1 generation? (consider the haplotypes that are transmitted by this F1 generation to the F2 generation).

No. The frequency of A_1B_1 gametes will be the same as A_1B_2, A_2B_1, and A_2B_2. (see problem 3). Therefore, there is no allelic association, which is when the frequency of alleles transmitted together is different from the frequency expected from the their independent frequencies. In this case, the frequency of A and B are both 0.5 and the frequency of A_1B_1 gametes is precisely (0.5)(0.5) = 0.25

5. (1 point) What is the expected frequency of each of the nine possible genotypes in the F2 progeny (assuming random mating among the F1)?

A 4 x 4 Punett square with the four equally likely haplotypes (A_1B_1, A_2B_1, A_1B_2 and A_2B_2) along each axis yields the following:

$A_{1,1}$ $B_{1,1}$	**1/16**
$A_{1,1}$ $B_{1,2}$	**2/16**
$A_{1,1}$ $B_{2,2}$	**1/16**
$A_{1,2}$ $B_{1,1}$	**2/16**
$A_{1,2}$ $B_{1,2}$	**4/16**
$A_{1,2}$ $B_{2,2}$	**2/16**
$A_{2,2}$ $B_{1,1}$	**1/16**
$A_{2,2}$ $B_{1,2}$	**2/16**
$A_{2,2}$ $B_{2,2}$	**1/16**

This is just a di-hybrid cross (as discussed in detail in section 3.2 of the text).

6. (1 point) Is locus A at Hardy-Weinberg equilibrium in the F2 generation (your answer would be the same for locus B)?

Yes, locus A is in Hardy-Weinberg equilibrium in this generation (as you would expect, since there was random mating). The two allele frequencies are p = 0.5 nd q =0.5.
The two types of homozygotes are each 1/4, as predicted.

$A_{1,1}$ **= 4/16 =** $\mathbf{p^2}$
$A_{1,2}$ **= 8/16 = 2pq** **(remember p + q = 1)**
$A_{2,2}$ **= 4/16 =** $\mathbf{q^2}$

The fact that the freqencies of the genotypes (1/4, 1/2 and 1/4) fit the Hardy-Weinberg equilibrium for p = 1/2, q = 1/2 is all that matters. A locus can be at Hardy-Weinberg equilibrium even if not all of the assumptions that go into deriving the Hardy-Weinberg equation hold.

Answer key

Later (questions 7-9), you allow a large and equivalent number of individuals from the two populations -- for example, 500 males and 500 females from population 1 and 500 males and 500 females from population 2 -- to mate at random (and they do mate at random).

7. (1 point) Is locus A at Hardy-Weinberg equilibrium in the "G1" generation?

Random mating means that half of the matings of population 1 males are with females from population 1 and half are with females from population 2 (and so on). To spell it out in detail, there are 1,000 matings.

250 will be pop 1 male by pop 1female (and all of the progeny will be $A_{1,1}$ $B_{1,1}$)
250 will be pop 1male by pop 2 female (and all of the progeny will be $A_{1,2}$ $B_{1,2}$)
250 will be pop 2 male by pop 1 female (and all of the progeny will be $A_{1,2}$ $B_{1,2}$)
250 will be pop 2 male by pop 2 female (and all of the progeny will be $A_{2,2}$ $B_{2,2}$)

Considering just the A locus, the genotype frequencies will be 1/4 $A_{1/1}$ 1/2 $A_{2,2}$ 1/4 $A_{2,2}$
So, **Yes, locus A is at Hardy-Weinberg equilibrium.**
Once again, the allele frequencies are p = q = 0.5 and the genotype frequencies fit that.

8. (1 point) What is the expected frequency of each of the four possible gametes transmitted from the G1 to the G2. The possible genotypes are A_1 B_1, A_1 B_2 , A_2 B_1 and A_2 B_2

The expected genotypes of the F1 generation are:

$A_{1,1}$ $B_{1,1}$	1/4 of population – generates all A_1B_1 haplotypes
$A_{1,2}$ $B_{1,2}$	1/2 of population – generates 1/4 A_1B_1 haplotypes, 1/4 A_1B_2 haplotypes, 1/4 A_2B_1 haplotypes, 1/4 A_2B_2 haplotypes
$A_{2,2}$ $B_{2,2}$	1/4 of population – generates all A_2B_2 haplotypes

The expected haplotypes are therefore:
3/8 A_1B_1 (1/4 from $A_{1,1}B_{1,1}$ + 1/8 from $A_{1,2}B_{1,2}$), **1/8 A_1B_2**, **1/8 A_2B_1**, **3/8 A_2B_2** (1/4 from $A_{2,2}B_{2,2}$ + 1/8 from $A_{1,2}B_{1,2}$).

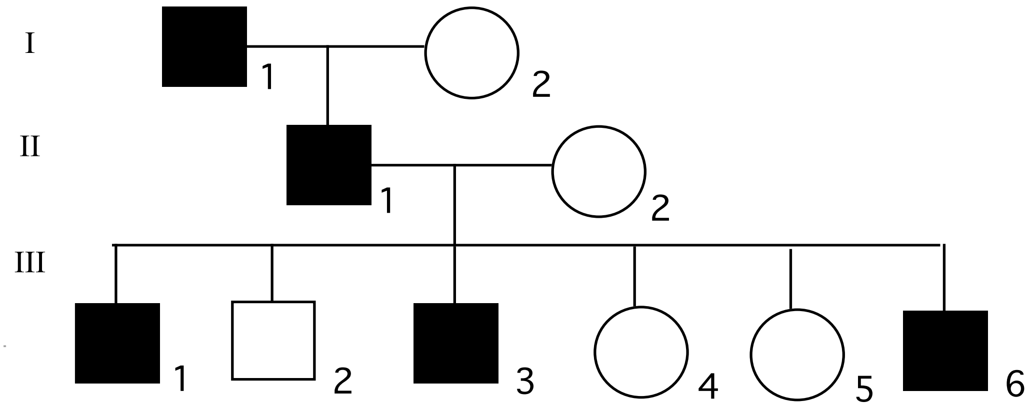
9. (1 point) Do the two alleles A_1 and B_1 show genetic association in this G1 generation?

Yes. The frequency of A_1 is 0.5 and the frequency of B_1 is 0.5 but the frequency of A_1 B_1 gametes is 3/8 rather than 1/4. Conversely, the frequency of A_1 is 0.5 and the frequency of B_2 is 0.5 but the frequency of A_1 B_2 gametes is 1/8 rather than 1/4.

This pedigree shows a family affected by an autosomal dominant genetic disease.

Genotypes for three linked markers, A, B and C, are shown

Answer key



The genotypes are:

I-1	$A_{1,2}$ $B_{1,2}$ $C_{1,2}$
I-2	$A_{3,3}$ $B_{3,3}$ $C_{3,3}$
II-1	$A_{1,3}$ $B_{1,3}$ $C_{1,3}$
II-2	$A_{4,4}$ $B_{4,4}C_{4,4}$
III-1	$A_{1,4}$ $B_{1,4}$ $C_{1,4}$
III-2	$A_{3,4}$ $B_{3,4}$ $C_{3,4}$
III-3	$A_{1,4}$ $B_{3,4}$ $C_{3,4}$
III-4	$A_{3,4}$ $B_{3,4}$ $C_{1,4}$
III-5	$A_{3,4}$ $B_{3,4}$ $C_{3,4}$
III-6	$A_{1,4}$ $B_{1,4}$ $C_{1,4}$

10. (1 point) Indicate the phase of alleles in individual II-1 by showing his haplotypes. There are four possibilities. They are

a) A_1 B_1 C_1 / A_3 B_3 C_3
b) A_1 B_1 C_3 / A_3 B_3 C_1
c) A_1 B_3 C_3 / A_3 B_1 C_1
d) A_1 B_3 C_1 / A_3 B_1 C_3

The answer is
a) $A_1B_1C_1/A_3B_3C_3$
since the A_1 B_1 C_1 alleles all came from the same parent (the father).

11. (1 point) What is the order of these three markers? (ABC, ACB or BAC)?

The order should be **ABC,** due to the fact that there is one A recombinant (III-3) and one C recombinant (III-4). Therefore, it is highly unlikely that either allele is in the middle due to the fact that a rare double recombination event would be required to produce the results observed.

Answer key

---

12. (2 points) Ignoring all of the other loci, calculate a lod score for linkage to of the disease to A with $\theta = 0$

There are 0 recombinants observed between A and the disease (everyone who has A_1 has the disease). Therefore, you can use the special lod score case Z = 0.3n, where n = the number of nonrecombinants (n = 6 in this case). **Z = 1.8**

$$Z = \log \frac{(1\text{-}0)^6}{(0.5)^6} = \log (2)^6 = 6 \log(2) = 6(.3) = 1.8$$

13. (2 points) Now, assume that this disease is only 75% penetrant. What is the lod score for linkage to A with $\theta = 0$ under this revised model?

In the given pedigree, individuals in generation III either get the marker A1 and the disease, or A3 and no disease. If the disease were 100% penetrant and linked to the marker, the probability of getting A1 and the disease is 0.5, just like the probability of getting A3 and not getting the disease (probabilities have to add up to 1).

In this case, however, 25% of the people with the marker will not get the disease, resulting in the following probabilities:

	$\theta = 0$ (complete linkage)	$\theta = 0.5$ (no linkage)
B1 transmitted, disease	0.375	0.1875
B1 transmitted, no disease	0.125	0.3125
B3 transmitted, disease	0	0.1875
B3 transmitted, no disease	0.5	0.3125

For example, the probability of individual III-1 turning out the way he did is 0.375, and the probability for individual III-2 is 0.5 **assuming linkage!**

Your null condition in this case would be to assume nonlinkage, or $\theta = 0.5$. Assuming nonlinkage and 100% penetrance, each of the conditions is equally likely (probability of 0.25). 75% penetrance would shift 25% of each diseased group to the nondiseased counterpart (25% of A1 transmitted, disease shifted to A1 transmitted, no disease) resulting in the probabilities seen in the second column above.

$$\text{Now, } Z = \log \frac{\text{(probability everyone turned out the way they did assuming linkage)}}{\text{(probability everyone turned out the way they did assuming nonlinkage)}}$$

$$\text{or } \quad = \log_{10} \frac{(0.375 * 0.5 * 0.375 * 0.5 * 0.5 * 0.375)}{(0.1875 * 0.3125 * 0.1875 * 0.3125 * 0.3125 * 0.1875)} = 1.515$$

14. (1 point) Ignoring all other loci, what value of θ would give the highest lod score for linkage of the disease to B?

There is 1 recombinant out of 6 with regards to B, III-3. Therefore, **θ = 1/6** (the observed recombination rate) will yield the highest lod score. You must show your work to receive full credit if you plugged in multiple values for θ to solve this problem.

15. (1 point) What is the value of that maximal lod score (for linkage of the disease to B at the value of θ that gives the highest possible lod score)?

$$Z = \log_{10} \frac{(5/6 * 5/6 * 5/6 * 1/6 * 5/6 * 5/6)}{(1/2 * 1/2 * 1/2 * 1/2 * 1/2 * 1/2)} = .632$$

16. (1 point) Ignoring all other loci, what value of θ would give the highest lod score for linkage of the disease to C?
17. (1 point) What is the value of that maximal lod score (for linkage of the disease to C at the value of θ that gives the highest possible lod score)?

See #14 and 15 for explanations, **θ = 2/6** (the observed recombination rate), **Z = 0.148**