

7.03 Fall 2003
Problem Set #7
Solutions Key

1. (a) Let's denote the disease allele as - and the wild-type allele as +. The "grandfather's" genotype (top row) at the disease locus must be +/-, since not all of his offspring (the "parents") are affected by the autosomal dominant disease. Therefore, all of the affected individuals on this pedigree are heterozygous (+/-) at the disease locus.

We can determine the phase relationship between the disease and SSR62 in each of the affected parents (second row) by looking at the genotypes of the grandparental generation and determining what alleles each grandparent passed on to affected parent. At SSR62, the affected mother received allele B from the unaffected grandmother. She also received a + at the disease locus from the grandmother, whose genotype at the disease locus is +/+. Therefore, the affected mother must have received the rest of her alleles, SSR62 A and the disease allele -, together from the affected grandfather. So we know the phase relationship in the affected mother looks like the following:

SSR62	Disease	
A	-	
B	+	

Similarly, the phase in the affected father can also be found by looking at what alleles he received from the grandparental generation. The phase in the affected father looks like the following:

SSR62	Disease	
A	-	
C	+	

Now we look at the offspring generation (bottom row). In all cases, SSR62 A is in phase with -, the disease allele. In the four offspring on the left side, the three affected offspring received SSR62 A and - from the affected mother. These three belong to the parental (non-crossover) class. But the unaffected daughter received SSR62 A and + from the affected mother, which only could have happened if there were a crossover event between the SSR62 locus and disease locus. Therefore, she belongs to the recombinant (crossover) class. On the right, the affected father gave his two affected sons SSR62 A and -, and gave SSR62 C and + to his two unaffected children, so all four children on the right are in the parental class.

Among the eight offspring, there **are seven parentals and one recombinant**, so the LOD score calculation looks like the following:

$$\text{LOD}_{\theta=0.1} = \log [(0.45)^7(0.05) / (0.25)^8] = 1.088$$

We are not finished yet. We must also consider the three parents in the middle row. However, we do not know the phase of the affected grandfather, so the LOD score calculation for the three middle row parents will involve the phase unknown formula. The two possible phases for the affected grandfather are:

SSR62	Disease	
A	-	Phase I
B	+	

SSR62	Disease	
B	-	Phase II
A	+	

In the middle row, the two affected parents in got SSR62 A from the affected grandfather, while the unaffected female got SSR62 B from the affected grandfather. So **if we assume phase I, there are three parentals. If we assume phase II, there are three recombinants.** The LOD score calculation would look like the following:

$$\text{LOD}_{\theta=0.1} = \log \{ [0.5(0.45)^3 + 0.5(0.05)^3] / (0.25)^3 \} = 0.465$$

The total LOD score for linkage between SSR62 and the disease in this family would be the sum of the two scores, since the events are independent of each other.

$$\text{LOD}_{\theta=0.1} = 1.088 + 0.465 = 1.553$$

(b) Using the same reasoning as in part a, we can determine the phase relationship between SSR93 and the disease in the affected mother and father:

Affected mother:

SSR93	Disease	
B	-	
A	+	

Affected father:

SSR93	Disease	
B	-	
A	+	

Now that we know the phase, we must look at the eight offspring. The only recombinant is the unaffected daughter on the right side, who received SSR93 B and + from the

affected father. All the other seven offspring are parentals, since they received from the affected parents either SSR93 B and - or SSR93 A and +.

Again, **there are seven parentals and one recombinant**. So the LOD score is the same as in part (a):

$$\text{LOD}_{\theta=0.1} = \log [(0.45)^7(0.05) / (0.25)^8] = 1.088$$

Again, we must also consider the three parents in the middle row. The two possible phases for the affected grandfather are:

SSR93	Disease	
B	-	Phase I
C	+	

SSR93	Disease	
C	-	Phase II
B	+	

In the middle row, the two affected parents both got SSR93 B from the affected grandfather, while the unaffected female got SSR93 C from the affected grandfather. So **if we assume phase I, there are three parentals. If we assume phase II, there are three recombinants**. The LOD score calculation would look like the following:

$$\text{LOD}_{\theta=0.1} = \log \{ [0.5(0.45)^3 + 0.5(0.05)^3] / (0.25)^3 \} = 0.465$$

The total LOD score for linkage between SSR93 and the disease in this family would be the sum of the two scores, since the events are independent of each other.

$$\text{LOD}_{\theta=0.1} = 1.088 + 0.465 = 1.553$$

(c) To calculate linkage between the two SSR's, we disregard whether a person is affected or unaffected and turn our attention exclusively to the SSR genotypes. We are no longer restricted to following the alleles of the two affected parents. Therefore, we can consider the passage of alleles from all four parents and the two grandparents to their respective offspring independently:

(i) Left side middle row father: This father is homozygous for SSR62 C. Therefore, **he is uninformative for the LOD score calculation**, since crossing over between the SSR62 and SSR93 loci will make no difference in the alleles that are passed on (i.e. we cannot distinguish between recombinant and parental classes).

(ii) Left side middle row mother: This mother is heterozygous at both SSR loci, so she is statistically informative. We can also determine her phase by looking at what alleles she received from the grandparents. From the grandmother, she definitely received SSR93 A and SSR62 B, so these two alleles are in phase. Therefore, she received her other alleles, SSR62 A and SSR93 B, together from the grandfather. Her phase looks like the following:

SSR62	SSR93
A	B
B	A

Looking at the four offspring on the left, we can see that they all received SSR62 C from their father and SSR62 A from their mother. At SSR93, three offspring received SSR93 B from their mother. They are parentals, since SSR62 A and SSR93 B are in phase. The unaffected daughter received SSR93 A and SSR62 A from her mother. This could only have happened if there were a crossover between the two SSR loci, so she is a recombinant. With **three parentals and one recombinant**, the LOD score calculation looks like the following:

$$\text{LOD}_{\theta=0.1} = \log [(0.45)^3(0.05) / (0.25)^4] = 0.067$$

(iii) Right side middle row father: This father received SSR93 A and SSR62 C from the grandmother, so these two alleles are in phase. Therefore, he received his other alleles, SSR62 A and SSR93 B, together from the grandfather. His phase looks like the following:

SSR62	SSR93
C	A
A	B

Looking at the four offspring on the right, we can see that there are again three parentals and one recombinant. The first two sons both received SSR62 A and SSR93 B from their father and the last son received SSR62 C and SSR93 A from the father. So these three are the parentals. The daughter received SSR62 C and SSR93 B from the father. She is a recombinant. With **three parentals and one recombinant**, the LOD score is the same as above in (ii):

$$\text{LOD}_{\theta=0.1} = \log [(0.45)^3(0.05) / (0.25)^4] = 0.067$$

(iv) Right side middle row mother: This mother is homozygous for SSR93 C, so she is uninformative for the same reason as the left side father.

(v) Top row grandfather: The phase relationship between the two SSRs is unknown in this individual. The two possible phases are:

SSR62	SSR93	
A	B	Phase I

B	C

SSR62	SSR93	
B	B	Phase II

A	C

In the middle row, the three parents (left to right) received the following SSRs from the grandfather: SSR62A SSR93B, SSR62B SSR93C, and SSR62A SSR93B. So **if we assume phase I, there are three parentals. If we assume phase II, there are three recombinants.** The LOD score calculation would look like the following:

$$\text{LOD}_{\theta=0.1} = \log \left\{ \frac{0.5(0.45)^3 + 0.5(0.05)^3}{(0.25)^3} \right\} = 0.465$$

(vi) Top row grandmother: She is uninformative because she is homozygous at the SSR93 locus.

We can add the LOD scores for the left side mother, right side father, and the grandfather because they passed on alleles to their offspring independently (for the same reason we can add the LOD scores from different families). So the final LOD score for linkage between SSR62 and SSR93 in this family is:

$$\text{LOD}_{\theta=0.1} = 0.067 + 0.067 + 0.465 = 0.599$$

(d) From now on, the alleles for SSR93 will be known as A', B', and C'. To determine the relative order of the three markers, we must look at the two individuals who were considered recombinants in parts (a) and (b). In part (a), the recombinant individual is the unaffected female in the last row (second from left). She received the following alleles from her affected mother: A, A', +. In the affected mother, the three possible allele orders are (phase known):

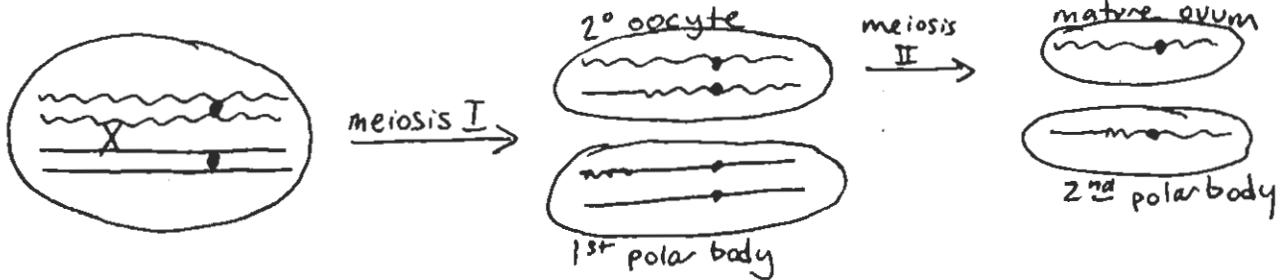
62	93	disease	
A	B'	-	order 1

B	A'	+

93	62	disease	
B'	A	-	order 2

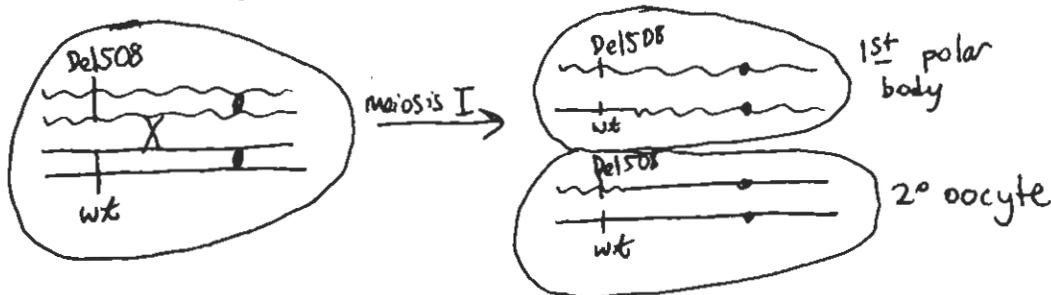
Note: the best solution to this question (the one described above) was very subtle and required a familiarity with material introduced earlier in the course, namely the superiority of three factor crosses in determining genetic order and distance. A common incorrect solution to this problem involved using the LOD scores calculated in earlier parts to predict order. This method is weaker because LOD scores essentially measure two factor distances (they are two factor crosses between two markers).

2) During ovulation, the primary oocyte divides into a polar body and a secondary oocyte. This division is the meiosis I division. As in any meiosis I division, it follows any recombination events, and it leads to the separation of homologous chromosomes via the centromeres. Next, the secondary oocyte undergoes meiosis II, which is the separation of sister chromatids, via the centromeres. This produces a second polar body and the mature ovum.

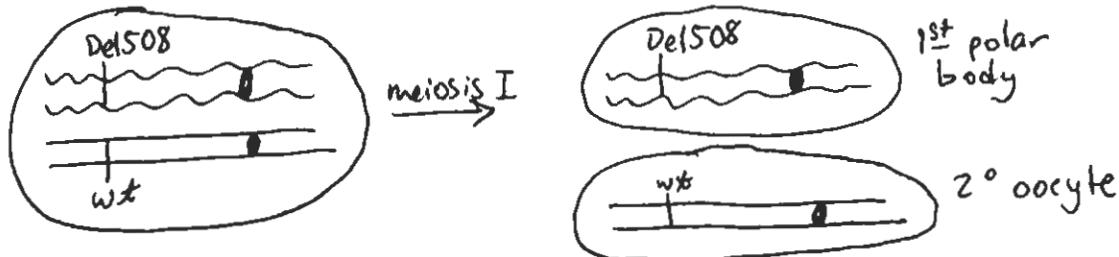


What we are observing in these data are the genotypes of the polar bodies. These polar bodies are being tested because they will not go on to become the egg (mature ovum) that would be fertilized by the sperm.

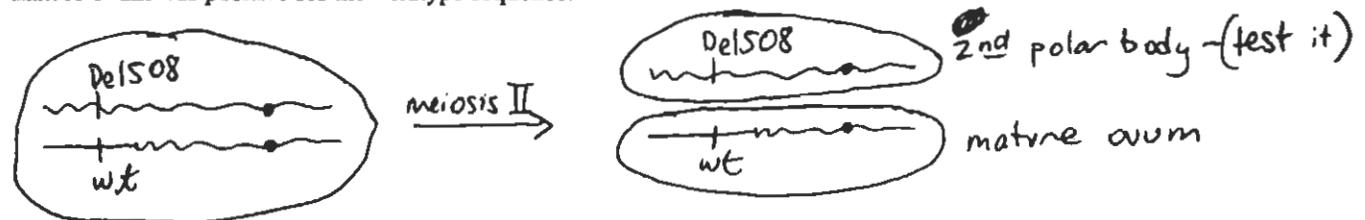
a) The polar bodies of oocytes #1 and #5 test positive for both the Del508 and the Wildtype sequence because a recombination event occurred between the centromere and the site of the cystic fibrosis gene prior to Prophase I. The other oocytes test positive for only one of the two sequences because no such recombination event occurred.



b) Since the secondary oocyte will go on through further development (eventually to make the mature ovum), and the polar body that we are testing will not develop further, we want the polar body to carry both copies of Del508. If the polar body carries both copies of Del508, we can assume that the secondary oocyte will have both of the wildtype copies. Therefore, we would choose oocytes #2 or #3 to use in *in-vitro* fertilization.



c) The Meiosis II division creates the mature ovum and a second polar body, from the secondary oocyte. If we could PCR test the second polar body and then return the mature ovum to the women's uterus, then we would choose an ovum where the second polar body tested positive for the Del508 (and the first polar body was already known to be positive for both sequences). If the second polar body of oocytes #1 or #5 tested positive for the Del508, then we would feel confident in returning the mature ovum from these oocytes to the women's uterus, knowing that the mature ovum was positive for the wildtype sequence.



d) If we were interested in identifying the parental origin of an additional chromosome 7, we could use polymorphisms in the cystic fibrosis gene, assuming each parent has a unique set of polymorphisms. Suppose the mother is A/B and the father is C/D. Then, if the child were A/B/C, we would know that the additional chromosome came from the mother. There are a number of other genotypes that would allow us to make the same conclusion (for instance, if the mother was A/A and the father C/C, we could still make an accurate conclusion).

If no recombination occurred, then we could indeed identify whether non-disjunction occurred in meiosis I or meiosis II (for instance, given parental genotypes A/B (mother) and C/D (father), the genotype A/B/C would indicate non-disjunction in the mother in meiosis I, whereas the genotype A/A/C would indicate non-disjunction in the mother in meiosis II. However, since we know already that recombination occurs with some frequency, we cannot here use the polymorphisms in the cystic fibrosis gene to make any conclusions about whether non-disjunction occurred in meiosis I or meiosis II, based on the child's genotype.