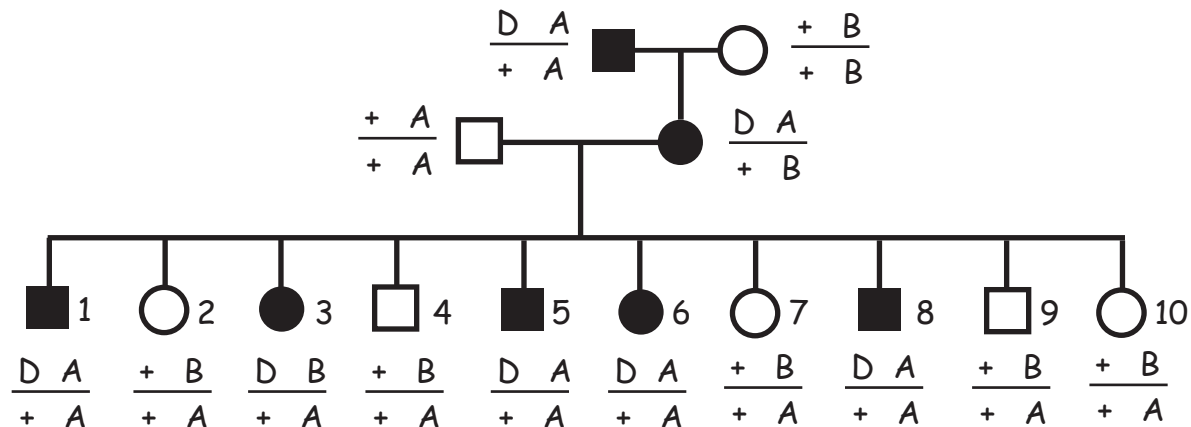


Lecture 11

Consider the following pedigree of a large family showing segregation of a dominant marker D and an SSR with alleles A and B.



The phase of the doubly heterozygous mother is known and most of the ten children appear to show cosegregation of the trait and the SSR allele indicating that this pedigree will be highly informative. However individual 3 appears to have received a crossover gamete. If we calculated the LOD score as before, $P(\text{data} \mid \text{completely linked}) = 0$ and the LOD would be $-\infty$. To calculate a LOD score for linkage at a distance $\neq 0$, we use the formula:

$$\text{LOD}_{\theta} = \log_{10} \frac{P(\text{observed segregation} \mid \text{loci linked by } \theta)}{P(\text{observed segregation} \mid \text{loci unlinked})}$$

θ = Recombination fraction (complete linkage, $\theta = 0$; unlinked, $\theta = 0.5$)

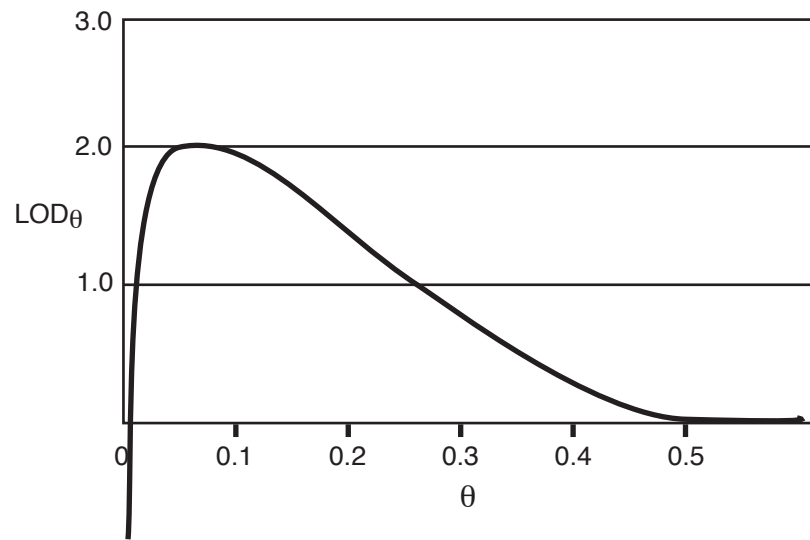
For this example, the best estimate for the recombination fraction is 0.1 because we see 1 crossover gamete and 9 parental gametes out of 10 total.

$$P(\text{data} \mid \text{loci linked by } 0.1) = (1/2 \cdot 0.9)^9 (1/2 \cdot 0.1)^1$$

$$P(\text{data} \mid \text{loci unlinked}) = (1/4)^{10}$$

$$\text{LOD}_{0.1} = \log_{10} \frac{(0.9^9 \cdot 0.1) (1/2)^{10}}{(1/4)^{10}} = \log_{10} 0.039 \cdot 2^{10} = 1.6$$

Imagine that multiple families that showed segregation of D and the and SSR were analyzed. To estimate the distance between these markers the total LOD score is calculated as a function of θ (this usually requires the computational power of a computer) to produce a plot that may look something like this:



In this example the maximum LOD score is at about $\theta = 0.08$. In other words the best estimate of the distance between D and the SSR is about 8 cM.