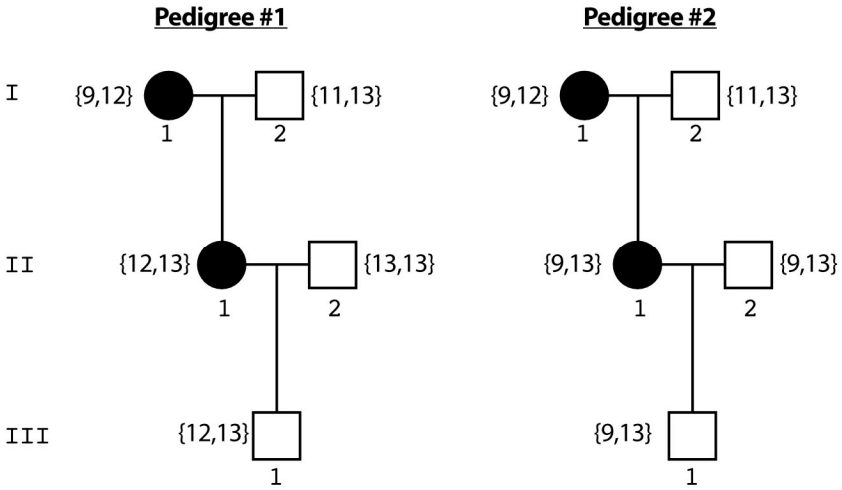
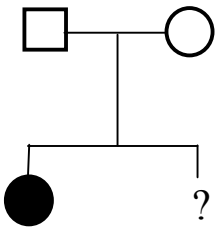
Problem Set 7
Genetics 371
Winter 2010

1. The two pedigrees below show inheritance of an autosomal dominant trait (D = disease, dominant; d = normal, recessive). Numbers in {curly brackets} indicate alleles of a microsatellite repeat polymorphic locus. For each pedigree, state whether the meiosis in II-1 is informative or uninformative, giving the parental types for II-1 in each case.



Pedigree 1: The meiosis is informative.
Pedigree 2: The meiosis is uninformative.

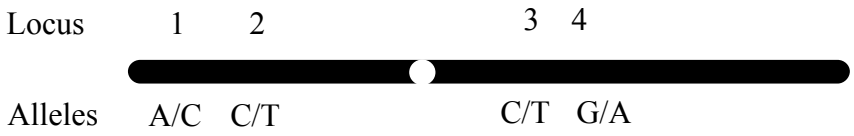
2. Consider the following pedigree that is segregating a recessive Mendelian disease:



If the population frequency of this disease is 10^{-5}, what is the sibling recurrent risk (λ_S)? If the frequency of this disease is 10^{-1} what is λ_S? Clearly state in one or two sentences why λ_S depends on the frequency of the disease in the general population.

If the frequency of the disease in the population is 10^{-5} the sibling recurrent risk is 25000. If the frequency of the disease is 10^{-1} the sibling recurrent risk is 2.5. Because the sibling recurrent risk is calculated by dividing by the population frequency, the risk of disease given that you have an affected sibling is higher when the frequency of disease in the population is rare (i.e., you are dividing by a smaller number).

3. Consider the diagram below indicating the locations and alleles at four loci along human chromosome 3:



(a) Is locus 1 more likely to show linkage with locus 2 or locus 3? Why?

Locus 1 is more likely to be linked to locus 2 because it is physically closer compared to locus 3.

(b) Are the alleles of locus 1 more likely to be associated with the alleles of locus 2 or 3? Why?

The alleles of locus 1 are more likely to be associated with the alleles of locus 2.

(c) Assume the recombination distance between locus 3 and 4 is 20 cM. What does this imply about patterns of association between the alleles of locus 3 and 4?

We cannot infer anything about whether the alleles between locus 3 and 4 show association even though we know that the two loci are linked. Remember that linkage is a property of loci (whether they separated by a map distance < 50 cM) whereas association is a property of alleles.

(d) Assume the recombination distance between locus 3 and 4 is 1 cM, but no association exists among their alleles. Explain in either words, or using a diagram, how this could be possible.

Although these two loci are very tightly linked, it is not surprising that the alleles are not associated. While recombination is infrequent between locus 3 and 4 (only 1% of gametes are recombinant) over many generations enough recombination occurs to remove any association.

4. Josh and Stan independently decide to perform an association study for Celiac's Disease. Each collects 1,000 cases and controls and genotype the same marker (which has alleles A and T). The contingency table from each study is shown below:

Josh's Study

	A	T	
Case	275	225	500
Control	225	275	500
	500	500	1000

Stan's Study

	A	T	
Case	290	210	500
Control	260	240	500
	550	450	1000

(a) Without doing a formal chi-square test, which study do you think will result in a more significant association and why?

Josh wins – the difference in allele frequencies between cases and controls is largest in his study.

(b) If Josh and Stan were studying the same disease and genotyped the same marker, explain how they could arrive at two different sets of numbers in the tables above.

There are several plausible explanations for why seemingly similar association studies give different results. The simplest reason is that one would expect to observe slightly different estimates of allele frequencies from one study to the next – this is called "sampling variation". Another valid explanation is that perhaps Josh and Stan performed there studies in two separate populations, and the frequency of either the SNP alleles for the marker tested above or the disease allele varies between populations.

(c) Assume the chi-square value for Josh's study is 9.6 leading to a p-value of 0.0019. What is the null hypothesis and how do you interpret the p-value? Do these data allow you to claim that the A allele causes Celiac's disease? Why or why not?

The null hypothesis is that there is no difference in allele frequency between cases and controls for the SNP marker. Because the p-value is less than 0.05, we reject the null hypothesis and conclude that this SNP is associated with Celiac's disease. However, we cannot claim that the A allele causes Celiac's disease because it could simply be associated (on the same haplotype background) with the true disease causing allele.

5. [*Before you panic: This question is too long to be an exam question, but we do expect you to understand and to be able to work any part of this kind of question.]*

Detailed answer provided in separate pdf.