

1) Read the review article: Sequencing technologies – the next generation (see website) Compare and contrast the strengths and weaknesses of the PacBio RS and the Illumina Hi-Seq. Give one good use case for both machines and one bad use case for both machines. Explain the logic behind your choices.

2) Computing a Log Odds (LOD) score can be useful for determining linkage of a SNP to a particular phenotype. In the progeny below there are 10 individuals' SNPs that are thought to be correlated with a particular disease gene (Disease/D and Normal/N)

Individual	1	2	3	4	5	6	7	8	9	10
SNP	'AA'	'AC'	'CC'	'AC'	'AA'	'AC'	'CC'	'AC'	'AA'	'AC'
	D	N	N	N	D	N	N	N	N	D

- What seems to be the SNP of the diseased and normal phenotypes?
- What is the probability of the observed phenotypes that the SNP is completely unrelated to the disease gene?
- What is the probability of the observed phenotype that there is a 10% of crossover to the SNP as compared to the disease gene?
- Compute the LOD score.
- Compute the LOD score for 20% and 30% recombination
- From the calculated LOD scores what is the optimal recombination fraction maximizing the LOD score?
- What is the penetrance of this SNP?

3) Genetics and Probability. Consider two loci 9 cM apart:

- What is the probability that there will be what is the probability that there will be recombination between them in one generation?
- What is the probability that there will not be recombination between them in one generation?
- What is the probability that there will NOT be recombination after 10, 20 and 30 generations?

- 4) Mendel's Peas. Suppose you are analyzing Mendel's data. He crossed a heterozygous plant RY/ry with a homozygous plant of genotype ry/ry. You want to see if he may have fudged his data, and if he didn't (we will assume he didn't) calculate the distance between the two alleles.

Phenotype	Count	Expected (Mendelian)
R/Y	546	?
R/y	52	?
r/Y	48	?
r/y	530	?

- What is the number that would correctly fill the right-hand column in the table above?
- What is the probability that these data fit the expected background distribution (hint: use a chi-square goodness of fit test)
- Calculate the recombination distance between the two loci.