

Name: \_\_\_\_\_

Section : \_\_\_\_\_

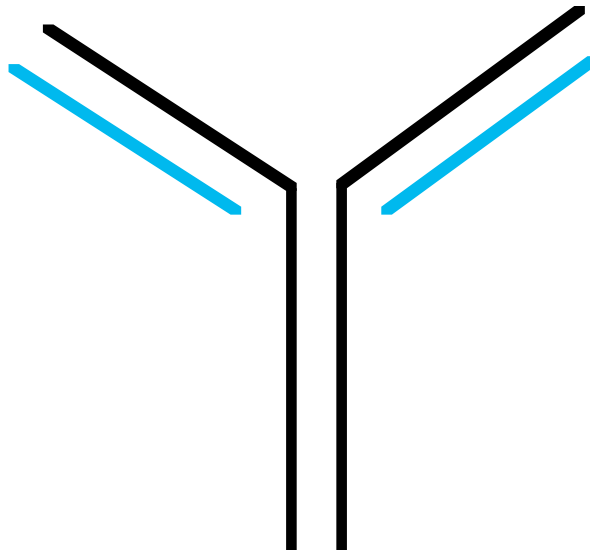
# 7.014 Problem Set 7

Answers to this problem set are to be turned in. **Problem sets will not be accepted late. Solutions will be posted on the web.**

## Question 1

### Part A

Shown below is a schematic representation of an antibody made in B cells.

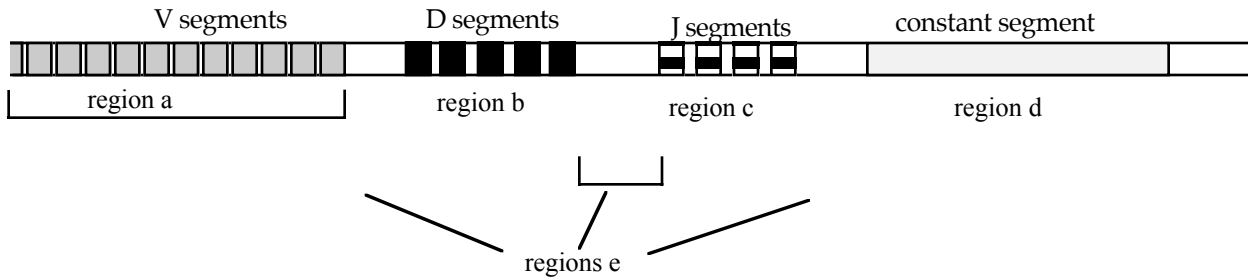


On the diagram label the following structures (a) – (d) on the antibody diagram.

- a) the antigen binding sites
- b) the heavy chains and the light chains
- c) the variable regions
- d)
  - i) If this was a membrane-bound antibody, indicate the membrane-spanning region of the antibody
  - ii) If this was a secreted antibody, indicate the region of the antibody that would interact with phagocytic cells
- e) List two ways that combinatorial joining generates antibody diversity.

## Question 1, continued

Below is a schematic of the unrearranged immunoglobulin heavy chain locus.



a) Indicate which region(s) contribute(s) to the formation of the antigen binding site of the antibody molecule.

b) After rearrangement, which region(s) would be the same in different B cells?

c) Which region(s) would never appear in any mRNA transcribed from this locus in a population of mature B cells?

d) Cindy wants to store cell samples from her chinchilla so that in the future she can produce a clone. She has isolated cell lines from the following four different cell types. However, freezer space in the lab is limited and she can keep only one. Which cell line should she keep?

Red blood cell      Gut cell      B cell      T cell

Why keep this cell line rather than the others?

## Question 1, continued

### Part B


The virus which causes AIDS (acquired immune deficiency syndrome) is called HIV (human immunodeficiency virus). HIV is a retrovirus, and its genome is a single RNA strand. Two copies of this RNA genome are packaged with two copies of the retroviral enzyme, reverse transcriptase, within a protein capsid.

a) For HIV infection to occur, a surface protein on the virus must interact with a surface protein on the cell. HIV specifically infects the helper T cells of the human immune system. Give one explanation for how HIV infection is specific for helper T cells.

b) Once infection has taken place, the RNA genome has to be made into double-stranded DNA. This process is mediated by reverse transcriptase. Once a double-stranded DNA copy of the HIV genome has been made, it is integrated into the host cell genome. The integration event is mediated by an enzyme, integrase.

i) What are the three steps required to produce double-stranded DNA from the single strand RNA genome?

ii) Given that these three steps are all performed by reverse transcriptase, what are three enzymatic activities that reverse transcriptase must possess?

iii) The central dogma states that:  DNA → RNA → Protein. Which one of the enzymatic activities possessed by reverse transcriptase goes against the central dogma? Explain your answer.

The drug most widely used to combat AIDS is Azidothymidine (AZT). AZT is very similar to thymidine, except that the 3' hydroxyl (OH) group on the deoxyribose ring has been replaced by an azido (N<sub>3</sub>) group.

c) Which process of the life cycle of the HIV do you think is inhibited by AZT?

d) What side effects, if any, do you expect from AZT treatment?

e) AZT-resistant forms of HIV have been isolated. These mutant viruses have mutations in one of the viral genes. In which viral gene do you think are the mutations most likely to be found? Why?

## Question 2

a) Below is a Cohort Life Table for a population of gray squirrels. Fill in the remaining spaces:

Gray Squirrel Cohort Life Table

$x$	$n_x$	$l_x$	$d_x$	$m_x$	$L_x$	$T_x$	$e_x$	$b_x$	$l_x b_x$
0	500	1	0.7	0.7	325	794	1.588	0	0
1		0.3	0.06	0.2	135	469	3.127	0.5	0.15
2	120		0.048	0.2	108	334	2.783	2	
3	96	0.192	0.038	0.2				2	0.348
4	77				69.5	139.5	1.811	2	0.308
5	62	0.124			46.5	70	1.129		0.124
6	31	0.062	0.046	0.75				0	0
7	8	0.016	0.016	1	4	4	0.5	0	0
8	0	0	0	0	0	0	0	0	0

b) When the table was made, some of the values were measured and others were calculated from those values. Which three values in the table must be measured in order to construct the rest of the table?

The gray squirrels have a high infant mortality rate ( $m_x$ ), but those that survive tend to not die of natural causes until age five to seven. In the meanwhile, those that die do so because of a natural predator of the gray squirrel, the Sports Utility Vehicle (SUV).

c) Given the table above, what is the replacement rate ( $R_0$ ) of the population? Is this population increasing, decreasing or staying the same?

d) A migration of SUVs into the area has increased the mortality rate ( $m_x$ ) of all squirrels between the ages of one and four years by 0.2.

i) Calculate the number of animals ( $n_x$ ) and the proportion of the original cohort surviving to each age ( $l_x$ ) for the years one through five given this new mortality rate. For any fractional values of  $n_x$  round to the nearest whole number.

ii) What is the new replacement rate for the population? Is this population increasing, decreasing or staying the same?

## Question 2, continued

- e) Given the original conditions, what is the generation time ( $G$ ) of the gray squirrels? (*Hint:  $G = (\sum l_x b_x x) / R_0$* )
- f) What is the intrinsic rate of increase ( $r$ ) of the gray squirrels? (*Hint:  $r = \ln R_0 / G$* )
- g) In the space below, draw a graph showing the growth of the cohort and its offspring as a function of time. Recall that  $N_t = N_0 e^{rt}$ . Let your time axis go from 0 to 25 years.
- h) How many gray squirrels are present after 25 years? Is this realistic? If so, explain why. If not, explain what we have not taken into account.
- i) If the carrying capacity ( $K$ ) equals 2,000, make a very rough sketch of the growth of the cohort and its offspring on the graph that you have made above. Be sure to indicate which curve belongs to part g and which belongs to part i. Assume instantaneous feedback.

## Question 2, continued

Survivorship curves reveal information about survival strategies of species. They are plots of the survival over time of individuals in a cohort (group of individuals in same population, born in same time period).

j) Draw three different types of survivorship curves (label your axes).

k) Use your drawing above to indicate which type of survivorship curve best applies to the cohorts given below.

i) Cohort of 20 sugar maple seedlings growing in the shadow of the parent, a tall sugar maple in mature forest.

ii) Cohort of 100 babies born at Mass General Hospital.

iii) Cohort of 1000 zooplankton in a lake containing a stable fish population which preys on zooplankton.

iv) Cohort of 100 corn plants planted and cultivated on a farm.

### Question 3

The Hollywood producers are getting wistful for good old times. A leading production company decides to film a script entitled "Squatting Aardvark, Spitting Llama".

Naturally, a production of this magnitude needs to be filmed on location – in this case the peaceful meadows and grasslands of the high Andes in central Peru. Since the director of the film knows nothing about the llamas to be used in the film, she hires you as a naturalist to accompany the film crew.

While studying the llamas in the Andes, you notice that they can be divided into three classes based on their spitting ability: Type A spits 1 m, type B spits 2.5 m, and (the most vicious of them all) type C spits 4 m. After doing some genetic tests, you determine that the spitting distance is determined by a single locus, which you name D (for distance).

The three types of llamas can be summarized as follows:

Type	Genotype	Phenotype (spitting distance)
A	dd	1 m
B	Dd	2.5 m
C	DD	4 m

You survey the local llama population and find that out of the 400 llamas, there are 64 type C llamas.

a) If the llama population is in Hardy-Weinberg equilibrium, then

i) What is the frequency of the d allele?

ii) What is the frequency of the D allele?

iii) How many of each type of llama do you expect to find out of the 400?

Type A:

Type B:

Type C: 64

-----

400 total

In reality, you find that the population distributes as follows:

Type A: 112

Type B: 224

Type C: 64

b) Which genotype is over-represented when compared to the predicted Hardy-Weinberg distribution? Explain the discrepancy between the actual numbers and your predicted values.

c) There are a number of ways in which the distribution of alleles in a population can differ from the ideal Hardy-Weinberg distribution. Match each of the following scenarios with the name that describes this effect. Then list one (of the 5 given in lecture) assumption of the Hardy-Weinberg model that is violated in this scenario.

A. On rare occasions, parents, both of which are homozygous recessive (dd), have an offspring born that is heterozygous (Dd).

B. An Argentinean production company, also filming a movie in Peru about llamas, imports its own population of llamas into the Andes.

C. Deciding that they're going to be making a sequel, the studio decides to bring 25 llamas back with them to California. When you look at this population a few generations later, you notice that it is in equilibrium, but with different phenotype distributions than the original Peruvian population.

D. For the film, you import the squatting aardvarks from near Lake Victoria. Unexpectedly, the squatting aardvark turns out to be a rather keen llama hunter. [This reeks havoc upon the filming schedule]. The squatting aardvark detects the llamas by the sound of their spitting. Llamas that spit farther make more noise when they spit.

	<b>Effect</b>	<b>Scenario</b>	<b>Assumption violated</b>
	Natural Selection		
	Mutation		
	Founder		
	Gene Flow		

d) A typical rate of mutation is 1 mutation in a particular gene per  $10^6$  sperm or egg cells. The effect of mutation on genotype distributions is very small. Why, then, are mutations considered to be so important for evolution?

e) Let's apply these concepts to a human population. Cystic Fibrosis is a recessive genetic disease of the lungs determined by a single genetic locus. About 1 in every 2500 human children is born with cystic fibrosis.

i) Of the approximately 100 students in 7.014, how many do you expect to be carriers of the Cystic Fibrosis mutation?

ii) Cystic Fibrosis is a lethal disease and affected individuals usually die prior to starting a family. Given this selective pressure, why does this disease still persist in the population?



## Question 4

Researchers isolated the XYZ protein from 12 different organisms (some extinct and some present now) and determined the amino acid sequences of a particular region from each. The sequences are listed below:

Organism	Sequence of XYZ
1	KSTSTDIKSREV□
2	KTTATEIKSREV□
3	KTTATDIKSKEV□
4	KSTAVEIKSKLF□
5	KTTATEIKSKLV□
6	KTLATDIKSREV□
7	KSTATDIKSREV□
8	KSTATEIKSKLV□
9	KTTATEIKSKEV□
10	KSTATDVKSREV□
11	KSTAVEIKSKLV□
12	KTTATDIKSREV□

a) Which amino acid(s) are invariant throughout this evolutionary tree? Why might have these amino acids been conserved in an actual organism?

b) Complementation across species.

i) Would you expect this protein from organism 1 to complement a deficiency in organism 11? Why or why not?

ii) Would you expect one of the glycolysis enzymes from organism 1 to complement a deficiency in organism 11? Why or why not?

c) Suppose you were able to get the DNA sequences of this protein from the twelve species. Would you expect the tree constructed using the DNA sequences to be more or less accurate than the tree constructed using the protein sequences? Why or why not?